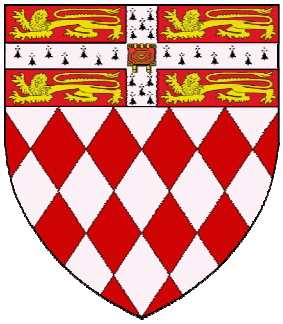
Structural and Biochemical Characterisation of the HR1 domains of the Protein Kinase C-related kinase family



Georgios Sophocleous

Fitzwilliam College

Department of Biochemistry

University of Cambridge

September 2018

A dissertation submitted for the degree of Doctor of Philosophy

Γηράσκω δ’ αἰεὶ πολλὰ διδασκόμενος.

As I grow old, I always learn many things

Solon the Athenian

Preface

This dissertation is a summary of research carried out in the Department of Biochemistry, University of Cambridge, UK, between October 2014 and September 2018. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the Degree Committee of the Faculty of Biology.

Acknowledgements

I would like to thank my supervisor, Dr. Helen Mott, for her invaluable help, advice, patience and guidance during the past four years. Her efforts were instrumental in my training as a scientist and for this I am very grateful.

I would also like to thank Dr. Darerca Owen for advice, experimental help and support.

I would particularly like to thank Professor Dame Jean Thomas for her support and inspiration.

I also thank past and present members of the DO/HRM group for their help as well as people of the Thomas and Nietlispach groups. I particularly owe gratitude to people who contributed at NMR Journal Club during the past few years for help with tackling NMR theory.

I would like to thank Dr. Katherine Stott who provided training with biophysical techniques and Dr. Peter Sharratt for amino acid analysis.

I would also like to thank the A.G. Leventis Foundation for providing partial funding of my PhD, as well as the Inner Wheel, Nicosia, Cyprus, Fitzwilliam College and the Philosophical Society, University of Cambridge who also made smaller but valuable contributions towards this project.

My thanks extend to Jesus College and particularly its second year undergraduates whom I supervised for the past four years; teaching them has taught me a lot.

I would also like to thank my friends and in particular Yi Lei, Daisy and Joyce for fun times as well as emotional support and encouragement. Studying here would not have been the same without them.

Above all I would like to thank my parents for their everlasting love, support and sacrifices to provide me with the best possible education, while always encouraging me to aim higher. Without them, the work of the past four years would not have been possible so this thesis is dedicated to them.

Name: Georgios Sophocleous

Title: Structural and Biochemical Characterisation of the HR1 domains of the Protein Kinase C-related kinase family

Summary

The Rho family of small GTPases act as molecular switches in the cell and are involved in signalling pathways implicated in cytoskeletal organisation. To mediate their effects, they interact with downstream effectors. The protein kinase C-related kinase (PRK) family of serine/threonine kinases consists of three members, PRK1, PRK2 and PRK3, all of which are Rho family effectors. PRKs have three N-terminal Homology Region 1 (HR1) domains, HR1a, HR1b and HR1c. These are predicted antiparallel coiled coils which interact with the Rho proteins. While HR1a and HR1b domains have been studied *in vitro*, very little is known about the HR1c domain. This project aims to provide much-needed biochemical, biophysical and structural understanding of HR1 domains.

The secondary structure and stability of PRK1 and PRK3 HR1 domains is outlined, followed by the identification of some HR1 domains as oligomerisation domains of their respective PRK. Analysis of the oligomeric state of the single, di- and tri-domains indicates that the three HR1 domains are not independent in solution. The NMR solution structure of the PRK1 HR1c domain is solved, highlighting the differences between this domain and other HR1 domains. The structure is then used as a tool to delineate the interactions between the three PRK1 HR1 domains. A model is created which summarises intra- and inter-molecular HR1 interactions. GTPase-HR1 interactions then expand the model by linking the HR1 oligomerisation to GTPase binding. New GTPase interactions with PRK1 and PRK3 HR1 domains are also identified, highlighting our limited knowledge about the PRKs and the subsequent studies which will need to follow.

Table of Contents

Preface i
Acknowledgements ii
Summary iii
Table of Contents iv
List of Abbreviations x
List of Figures xiv
List of Tables xvii
Chapter 1 Introduction 1
1.1 The Ras superfamily 1
1.1.1 The G domain 2
1.1.2 The switch regions 5
1.1.3 GEFs 8
1.1.4 GAPs 11
1.1.5 GDIs 14
1.1.6 Effectors 17
1.1.7 Post-translational modifications 18
1.1.8 Structural similarities and differences between the 5 families 21
1.2 The Rho family 23
1.2.1 Implications in several signalling pathways and cellular localisation 23
1.2.2 The three key players 25
1.2.3 Role in disease 25
1.3 Rho family effectors 27
1.3.1 Rac1 and Cdc42 effectors 27
1.3.2 RhoA effectors 30
1.4 The Protein Kinase C-related kinase family 34
1.4.1 Discovery 34
1.4.2 Domain arrangement 34
1.4.3 Activation 37
1.4.4 Autoregulation 40

1.4.5 Role in cellular signalling and physiological roles 41
1.4.6 Cellular localisation and tissue-specific expression 51
1.4.7 Role in disease 53
1.4.8 HR1 domains as Rho family effector domains 58
1.5 PRKs - The uncertain and the unknown 66
1.5.1 The uncertain 66
1.5.2 The unknown 67
1.6 Aims and experimental techniques 69
1.6.1 Aims 69
1.6.2 Experimental techniques 70
Chapter 2 Materials and methods 71
2.1 Materials 71
2.1.1 Chemicals 71
2.1.2 Commercial kits and other materials 71
2.1.3 Commercial enzymes 72
2.1.4 Oligonucleotides 72
2.1.5 Microbiological media 72
2.1.6 Buffers 73
2.1.7 Plasmids 75
2.1.8 Bacterial strains 77
2.1.9 Primers 78
2.2 Methods 79
2.2.1 Molecular cloning 79
2.2.2 Site-directed mutagenesis 83
2.2.3 SDS-PAGE 83
2.2.4 Expression trials 84
2.2.5 Protein expression 84
2.2.6 Protein purification 85
2.2.7 Determination of protein concentration 88
2.2.8 Mass spectrometry 89
2.2.9 Nucleotide exchange 89
2.2.10 Circular dichroism 90
2.2.11 Analytical ultracentrifugation 91
2.2.12 NMR Spectroscopy 92

2.2.13 Crystallisation trials 97
2.2.14 Isothermal Titration Calorimetry 98
2.2.15 Cross-linking 99
2.2.16 Pull-downs 100
2.2.17 Scintillation Proximity Assay 100
2.2.18 Computational methods 101
Chapter 3 Purification of PRK1 and PRK3 HR1 constructs and investigation of their oligomeric state 103
3.1 Introduction 103
3.2 Protein purification 103
3.2.1 Construct cloning 103
3.2.2 Expression trials 104
3.2.3 Purification of GST HR1 fusions using glutathione agarose beads 109
3.2.4 Removal of GST tag by proteolytic cleavage 109
3.2.5 Further purification of PRK1 HR1b, PRK1 HR1c, PRK1 HR1abc and PRK3 HR1c by gel filtration 110
3.2.6 Further purification of PRK1 HR1ab, PRK3 HR1a and PRK3 HR1b on a glutathione sepharose column followed by gel filtration 114
3.2.7 Further purification of PRK1 HR1a on a glutathione sepharose column followed by anion exchange chromatography and gel filtration 118
3.2.8 Further purification of PRK3 HR1ab and PRK3 HR1abc by cation exchange chromatograpy and gel filtration 120
3.2.9 Yields of purified proteins 124
3.3 Investigating the secondary structure of HR1 domains by circular dichroism 125
3.3.1 Introduction 125
3.3.2 Theory 125
3.3.3 Results 127
3.4 Investigating the oligomeric state of HR1 domains by analytical ultracentrifugation 133
3.4.1 Theory 133
3.4.2 Sedimentation velocity vs. sedimentation equilibrium 141
3.4.3 Absorbance vs. interference optics 142
3.4.4 Experimental conditions used 143
3.4.5 PRK1 HR1a 144
3.4.6 PRK1 HR1b 146
3.4.7 PRK1 HR1c 148

3.4.8 PRK1 HR1ab 150
3.4.9 PRK1 HR1abc 152
3.4.10 PRK3 HR1a 155
3.4.11 PRK3 HR1b 157
3.4.12 PRK3 HR1c 159
3.4.13 PRK3 HR1ab 161
3.4.14 PRK3 HR1abc 163
3.5 Discussion and conclusions 165
3.5.1 Conclusions about PRK1 oligomerisation 165
3.5.2 Conclusions about PRK3 oligomerisation 166
3.5.3 PRK1 and PRK3 oligomerisation comparison 166
3.5.4 Importance of oligomerisation 168
Chapter 4 Structural studies on the PRK1 HR1c domain 170
4.1 Introduction 170
4.2 Nuclear Magnetic Resonance 171
4.2.1 The NMR signal 171
4.2.2 The HSQC pulse sequence 180
4.3 NMR studies on PRK1 HR1c 184
4.3.1 2D-NOESY and 2D-TOCSY of unlabelled PRK1 HR1c 184
4.3.2 Purification of ^{15}N-labelled NMR sample 186
4.3.3 ^{15}N-HSQC and ^{13}C-HSQC experiments 188
4.3.4 Three-dimensional experiments for backbone NH group assignment 191
4.3.5 ^{15}N-separated TOCSY and ^{15}N-separated NOESY for backbone assignment 195
4.3.6 HCCH-TOCSY and ^{13}C-separated NOESY 198
4.3.7 Assignment of aliphatic side-chains 199
4.3.8 NH_2 group assignment 202
4.3.9 Assignment of Phe^{222} 202
4.3.10 Assignment of prolines 205
4.3.11 Medium and long range NOEs 205
4.3.12 Ambiguous and unambiguous NOEs 205
4.4 Structure calculations using ARIA 209
4.4.1 Dihedral angles 209
4.4.2 Distance restraints 211
4.4.3 Initial structure calculations 212

4.4.4 Manual refinement 212
4.4.5 The structure 215
4.4.6 Comparison with other HR1 domain structures 220
4.5 Relaxation analysis of the HR1c structure 223
4.5.1 Theory 223
4.5.2 ^{15}N-relaxation experiments on PRK1 HR1c 232
4.6 Investigating the structure of PRK3 HR1c 236
4.6.1 Sequence alignment of the two HR1c domains 236
4.6.2 Investigating the structure of PRK3 HR1c by NMR 237
4.7 Conclusions 239
Chapter 5 Investigating the intra- and inter-molecular interactions between HR1 domains 240
5.1 Introduction 240
5.2 Investigating the structure of the HR1abc tri-domain 240
5.2.1 Investigating HR1abc by NMR 241
5.2.2 Crystallisation trials 243
5.3 Investigating the interactions between the individual HR1 domains 245
5.3.1 Rationale 245
5.3.2 Isothermal Titration Calorimetry experiments 245
5.3.3 NMR titrations with individual domains 246
5.3.4 ^{15}N PRK1 HR1ab 260
5.3.5 Cross-linking experiments 267
5.4 Conclusions 273
Chapter 6 GTPase interactions with HR1 domains 276
6.1 Introduction 276
6.1.1 Known interactions 276
6.1.2 Aims 277
6.2 Investigating new interaction partners between the Rho family and the HR1 domains 278
6.2.1 Choice of GTPases to test 278
6.2.2 Choosing the best expression conditions 279
6.2.3 Pull-downs 281
6.2.4 Expression and purification of GTPases 284
6.2.5 Nucleotide exchange 285
6.2.6 Measuring the affinity of Cdc42 for PRK3 HR1c 286
6.2.7 Sequence alignments 288

6.3 How does oligomerisation affect GTPase binding? 290
6.3.1 RhoA binds to monomeric HR1a 290
6.3.2 RhoA binds to dimeric HR1a 290
6.3.3 Measuring the affinity of dimeric PRK1 HR1a for RhoA 293
6.4 Investigating the binding of RhoA to the HR1abc tri-domain 295
6.4.1 2-way pull-downs between G proteins and HR1abc 295
6.4.2 Measuring the affinity of PRK1 HR1abc for RhoA 297
6.5 Conclusions 299
Chapter 7 Discussion and future directions 304
7.1 Structural and biochemical insight into the HR1 domains 305
7.1.1 Intra- and inter-molecular interactions 305
7.1.2 Wider implications for PRK signalling in cells 307
7.2 Future work 310
7.2.1 Characterisation of the dimerisation interface 310
7.2.2 Characterisation of the HR1abc tri-domain 311
7.2.3 PRKs as Rho GTPase effectors 312
7.2.4 PRKs on the membrane 315
7.3 Concluding remarks 316
Appendix 1 317
Appendix 2 320
Appendix 3 322
Appendix 4 325
References 328

List of Abbreviations

ACK	Activated Cdc42-associated kinase
AR	Androgen receptor
ARIA	Ambiguous restraints for iterative assignment
AUC	Analytical ultracentrifugation
CD	Circular dichroism
Cdc42	Cell division control protein 42
CDK	Cyclin-dependent kinase
CIK	Citron kinase
CIP4	Cdc42-interacting protein 4
CRIB	Cdc42, Rac interactive binding
CV	Column volume
DH	Dbl homology
DHR1/2	Dock homology region 1/2
DMP	Dimethyl pimelimidate
DMS	Dimethyl suberimidate
DOCK	Dedicator of cytokinesis
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EGFR	Epidermal growth factor receptor
FBP17	Formin-binding protein 17
FTase	Farnesyl transferase
GAP	GTPase activating protein
GDI	Guanine nucleotide dissociation inhibitor
GDP	Guanosine diphosphate

GEF	Guanine nucleotide exchance factor
GGTase	Geranylgeranyl transferase
G_{α}	Heterotrimeric G protein α
G_{q}	Heterotrimeric G protein q
Graf	GAP for Rho associated with focal adhesion kinase
GTP	Guanosine triphosphate
H/D exchange	Hydrogen/deuterium exchange
HDAC	Histone deacetylase
HIDS	Hyperimmunoglobulinemia D syndrome
HM	Hydrophobic motif
HPV	Human papilloma virus
HR1	Homology Region 1
HSQC	Heteronuclear Single Quantum Coherence
HUVEC	Human umbilical endothelial cell
ICMT	Isoprenylcysteine carboxymethyltransferase
IL	Interleukin
INEPT	Insensitive nuclei enhanced by polarisation transfer
IQGAP	IQ-domain GTPase-activating protein
ITC	Isothermal titration calorimetry
LMP agarose	Low Melting Point agarose
LRR	Leucine-rich repeat
MAPK	Mitogen-activated protein kinase
mDia1	mammalian homolog of Drosophila diaphanous 1
MEFV	Mediterranean Familial Fever
MS	Mass spectrometry

mTOR	mammalian target of rapamycin
N-WASP	Neuronal Wiskott-Aldrich Syndrome protein
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
PAK	p21-activated kinase
PAR1	Protease-activated receptor 1
PDEδ	Phosphodiesterase δ
PDK1	Phosphoinositide-dependent protein kinase-1
PH	Pleckstrin homology
PI3-K	Phosphoinositide-3-kinase
PIF	PDK1-interacting fragment
PIP_2	Phosphatidylinositol 4,5-bisphosphate
PIP_3	Phosphatidylinositol 3,4,5-triphosphate
PK	Protein kinase
PKC	Protein kinase C
PBMC	Peripheral blood mononuclear cell
PRK	Protein kinase C-related kinase
Rac1	Ras-related C3 botulinum toxin substrate 1
Ras	Rat sarcoma
RBD	Rho-binding domain
RCE1	Ras-converting CAAX endopeptidase 1
Rho	Ras homolog
ROCK	Rho-associated coiled-coil kinase
SAXS	Small angle X-ray scattering

SE	Sedimentation equilibrium
SEC-SAXS	Size exclusion chromatography SAXS
SPA	Scintillation proximity assay
SV	Sedimentation velocity
TALOS	Torsion angle likeliness obtained from shift and sequence similarity
TAM	Tumour-associated macrophage
TM	Turn motif
Toca1	Transducer of Cdc42-dependent actin assembly protein 1
TOCSY	Total Correlation Spectroscopy
TPα	Thromboxane A_2 receptor α isoform
TPβ	Thromboxane A_2 receptor β isoform
TXA_2	Thromboxane A_2
WASP	Wiskott-Aldrich Syndrome protein

List of Figures

Figure 1.1 G proteins act as molecular switches in signalling pathways. 2
Figure 1.2 The G domain. 4
Figure 1.3 Structural features of the G domain. 7
Figure 1.4 Structure and function of GEFs. 10
Figure 1.5 Structure and function of GAPs. 13
Figure 1.6 Structure of Rac1 in complex with RhoGDI-1. 16
Figure 1.7 Prenylation of GTPases. 20
Figure 1.8 The 5 families of the Ras superfamily. 22
Figure 1.9 Phylogenetic tree of the Rho family GTPases. 24
Figure 1.10 Cytoskeletal structures controlled by Rho GTPases. 26
Figure 1.11 Rho GTPase/effector complex structures. 29
Figure 1.12 Summary of the key effectors of RhoA. 30
Figure 1.13 Domain structure of PRK proteins. 35
Figure 1.14 Activation of PRKs and downstream cellular effects. 38
Figure 1.15 Structure of PRK1 HR1b in complex with the *Salmonella* protein SspH1. 57
Figure 1.16 Sequence alignments of the HR1 domains of the three PRKs. 61
Figure 1.17 Structure of PRK1 HR1a interacting with RhoA. 64
Figure 1.18 Structure of PRK1 HR1b interacting with Rac1. 65
Figure 3.1 SDS-PAGE Analysis of PRK1 construct expression trials. 105
Figure 3.2 SDS-PAGE Analysis of PRK2 construct expression trials. 106
Figure 3.3 SDS-PAGE Analysis of PRK3 construct expression trials. 107
Figure 3.4 SDS-PAGE analysis of recombinant PRK1 HR1c protein expression and purification. 109
Figure 3.5 Further purification of PRK1 HR1b and corresponding SDS-PAGE analysis. 110
Figure 3.6 Further purification of PRK1 HR1c and corresponding SDS-PAGE analysis. 111
Figure 3.7 Further purification of PRK1 HR1abc and corresponding SDS-PAGE analysis. 112
Figure 3.8 Further purification of PRK3 HR1c and corresponding SDS-PAGE analysis. 113
Figure 3.9 Further purification of PRK1 HR1ab and corresponding SDS-PAGE analysis. 115
Figure 3.10 Further purification of PRK3 HR1a and corresponding SDS-PAGE analysis. 116
Figure 3.11 Further purification of PRK3 HR1b and corresponding SDS-PAGE analysis. 117
Figure 3.12 Further purification of PRK1 HR1a and corresponding SDS-PAGE analysis. 119
Figure 3.13 Further purification of PRK3 HR1ab and corresponding SDS-PAGE analysis. 122
Figure 3.14 Further purification of PRK3 HR1abc and corresponding SDS-PAGE analysis. 123
Figure 3.15 Circular dichroism principles. 126
Figure 3.16 PRK1 construct circular dichroism wavelength scans. 130
Figure 3.17 PRK3 construct circular dichroism wavelength scans. 131
Figure 3.18 Thermal stability comparison between PRK1 HR1c and PRK3 HR1c. 132
Figure 3.19 Analytical ultracentrifugation setup and principles. 136
Figure 3.20 PRK1 HR1a sedimentation velocity experiments. 145
Figure 3.21 PRK1 HR1b sedimentation velocity experiments. 147
Figure 3.22 PRK1 HR1c sedimentation velocity experiments. 149
Figure 3.23 PRK1 HR1ab sedimentation velocity experiments. 151
Figure 3.24 PRK1 HR1abc sedimentation velocity experiments. 153
Figure 3.25 PRK3 HR1a sedimentation velocity experiments. 156

Figure 3.26 PRK3 HR1b sedimentation velocity experiments. 158
Figure 3.27 PRK3 HR1c sedimentation velocity experiments. 160
Figure 3.28 PRK3 HR1ab sedimentation velocity experiment. 162
Figure 3.29 PRK3 HR1abc sedimentation velocity experiments. 164
Figure 3.30 Comparison of biophysical properties of HR1 domains. 168
Figure 4.1 The NMR signal. 173
Figure 4.2 The rotating frame and the effective field. 174
Figure 4.3 Pulse-acquire experiment. 176
Figure 4.4 Product operator approach and the HSQC pulse sequence. 182
Figure 4.5 2D NOESY and 2D TOCSY spectra of PRK1 HR1c. 185
Figure 4.6 SDS-PAGE Analysis of ^{15}N PRK1 HR1c expression trials in minimal medium. 186
Figure 4.7 Purification of ^{15}N PRK1 HR1c and corresponding SDS-PAGE analysis. 187
Figure 4.8 ^{15}N-HSQC spectrum of PRK1 HR1c. 189
Figure 4.9 ^{13}C-HSQC spectrum of PRK1 HR1c. 190
Figure 4.10 Triple resonance experiment magnetisation transfer pathways. 193
Figure 4.11 Backbone NH assignment using triple resonance experiments. 194
Figure 4.12 ^{15}N-separated NOESY and ^{15}N-separated TOCSY experiments. 195
Figure 4.13 ^{15}N-separated NOESY spectrum of PRK1 HR1c. 197
Figure 4.14 Magnetisation transfer pathways of HCCH-TOCSY and ^{13}C-separated NOESY. 198
Figure 4.15 Assignment of CH groups in a short aliphatic side-chain. 200
Figure 4.16 Assignment of CH groups in a longer aliphatic side-chain. 201
Figure 4.17 Assignment of side-chain NH_2 groups. 203
Figure 4.18 Assignment of Phe222 side-chain protons using CH-CH NOEs. 204
Figure 4.19 Assignment of proline side-chains. 207
Figure 4.20 Assignment of α-helical *i* to *i+3* NOEs. 208
Figure 4.21 The NMR solution structure of PRK1 HR1c. 216
Figure 4.22 Inter-helical contacts in the coiled coil. 217
Figure 4.23 The key long-range NOEs in the triple helical region. 218
Figure 4.24 The loop region NOEs. 219
Figure 4.25 Structural comparison of HR1 domains. 222
Figure 4.26 Correlation function plot. 225
Figure 4.27 Spectral density function plots. 226
Figure 4.28 Spectral density function at Larmor frequency. 227
Figure 4.29 ^{15}N relaxation measurements for PRK1 HR1c. 234
Figure 4.30 Sequence alignment of PRK1 HR1c and PRK3 HR1c. 237
Figure 4.31 2D NOESY and 2D TOCSY spectra of PRK3 HR1c. 238
Figure 5.1 2D NOESY titration spectra of PRK1 HR1abc. 242
Figure 5.2 PRK1 HR1abc crystallisation trials. 243
Figure 5.3 Chemical exchange regimes in NMR. 248
Figure 5.4 Quantitative analysis of ^{15}N HSQC spectra when HR1a was titrated into ^{15}N HR1c. 251
Figure 5.5 Quantitative analysis of ^{15}N HSQC spectra when HR1b was titrated into ^{15}N HR1c. 253
Figure 5.6 Quantitative analysis of ^{15}N HSQC spectra when HR1b was titrated into ^{15}N HR1c + HR1a. 255
Figure 5.7 Quantitative analysis of ^{15}N HSQC spectra when HR1a was titrated into ^{15}N HR1c + HR1b. 257
Figure 5.8 Quantitative analysis of ^{15}N HSQC spectra when HR1ab was titrated into ^{15}N HR1c. 259
Figure 5.9 ^{15}N-HSQC spectrum of PRK1 HR1ab. 262
Figure 5.10 Effect of NaCl on quality of the ^{15}N-HSQC spectrum of PRK1 HR1ab. 263
Figure 5.11 Comparison of PRK1 HR1b and PRK1 HR1ab. 264

Figure 5.12 Quantitative analysis of ^{15}N HSQC spectra when HR1c was titrated into ^{15}N HR1ab. 266
Figure 5.13 Chemistry of EDC and DMS cross-linking agents. 268
Figure 5.14 PRK1 HR1a cross-linking trials. 269
Figure 5.15 PRK1 HR1ab cross-linking trials. 270
Figure 5.16 PRK1 HR1abc cross-linking trials. 271
Figure 5.17 Cross-linking trials of PRK1 HR1ab and PRK1 HR1c. 272
Figure 5.18 Sequence alignment of PRK1 HR1 domains. 275
Figure 6.1 SDS-PAGE Analysis of Rho GTPase expression trials. 280
Figure 6.2 RhoA interacts with PRK1. 282
Figure 6.3 RhoD interacts with PRK1. 283
Figure 6.4 Purification of RhoA Δ4 F25N Q63L and corresponding SDS-PAGE analysis. 284
Figure 6.5 GMPPNP nucleotide exchange of GTPases. 285
Figure 6.6 Competition SPA to measure the affinity of Cdc42 for PRK3 HR1c. 287
Figure 6.7 Sequence alignment of RhoA and RhoD. 288
Figure 6.8 Sequence alignments of Rho family GTPases and their HR1 effector domains. 289
Figure 6.9 RhoA Δ4 F25N Q63L sedimentation velocity experiments. 292
Figure 6.10 Measuring the affinity of RhoA Δ4 F25N Q63L for the PRK1 HR1a dimer. 294
Figure 6.11 RhoA interacts with PRK1 HR1abc and PRK3 HR1abc. 296
Figure 6.12 Competition SPA to measure the affinity of RhoA for PRK1 HR1abc. 298
Figure 6.13 Summary of known PRK1 and PRK3 HR1 domain interactions with Rho GTPases. 299
Figure 6.14 Sequence alignment of Rho GTPases. 303
Figure 7.1 Model of RhoA binding to PRK1 HR1abc. 306
Figure 7.2 PRK1 activation and localisation pathways. 309
Figure 7.3 Possible supercomplexes between PRK1 HR1 domains and Rho family GTPases. 314
Figure A2.1 PRK1 HR1c MALDI analysis. 320
Figure A2.2 PRK3 HR1c MALDI analysis. 321

List of Tables

Table 1.1 The G1-G5 conserved sequence motifs. 3
Table 1.2 Key substrates of the PRKs. 41
Table 1.3 Sequence identity and similarity between the HR1 domains. 60
Table 2.1 Microbiological media. 72
Table 2.2 Cloning and transformation buffers. 73
Table 2.3 Protein purification buffers. 73
Table 2.4 Biochemical and biophysical analysis buffers. 74
Table 2.5 SDS-PAGE buffers. 75
Table 2.6 DNA plasmid vectors. 75
Table 2.7 Table of constructs. 76
Table 2.8 Bacterial strains. 77
Table 2.9 Table of DNA primers. 78
Table 2.10 PCR cycling program. 80
Table 2.11 Chromatography columns. 86
Table 2.12 Purification and biophysical analysis buffers for each protein. 87
Table 2.13 NMR experiments. 93
Table 2.14 Custom crystallisation plates. 98
Table 2.15 Cross-linking reactions with EDC. 99
Table 2.16 Cross-linking reactions with DMS/DMP. 99
Table 2.17 Software. 101
Table 3.1 Summary of expression trials. 108
Table 3.2 Final yields of purified proteins obtained. 124
Table 3.3 Summary of CD data. 128
Table 4.1 Experimental restraints and structural statistics. 214
Table 6.1 Summary of published G protein/PRK1 HR1 domain affinity data. 277
Table A3.1 Summary of AUC results. 322
Table A4.1 PRK1 HR1c chemical shifts. 327

Chapter 1 Introduction

1.1 The Ras superfamily

The Ras superfamily of small GTPases or G proteins consists of over 160 proteins which are often described as "molecular switches" of signalling pathways, based on their ability to exist in two states. G proteins bind to either guanosine diphosphate (GDP) or guanosine triphosphate (GTP). They can also hydrolyse GTP and it is this activity which classifies them as switches, as effectors only interact with the GTP-bound state, i.e. in a nucleotide-dependent manner. While the molecular switch mechanism is universally shared across the Ras superfamily, the range of cellular processes which are governed by this is astonishingly diverse. These processes include gene expression, vesicular trafficking, cytoskeletal organisation and nuclear import and export, amongst others.

The Ras superfamily can be divided into five distinct families on the basis of structural and functional similarities. These are the Ras, Rho (Ras homology), Rab, Arf and Ran families (reviewed in Wenneburg *et al.*, 2005). The proteins in this family all share a core G domain which is ~20 kDa in size (reviewed in Bourne *et al.*, 1991) and typically binds GDP and GTP with pM affinities. The cycling from the GDP- to the GTP-bound state and vice-versa depends on a number of regulator proteins (Figure 1.1). Guanine nucleotide exchange factors (GEFs) catalyse the exchange of GDP for GTP, activating the GTPase. Conversely, GTPase activating proteins (GAPs) stimulate the intrinsic hydrolysis activity of the GTPase to switch it off, thereby terminating its ability to interact with downstream effectors. Guanine nucleotide dissociation inhibitors (GDIs) provide an additional layer of regulation for the Rho and Rab families by extracting the GDP-bound GTPases from membranes and sequestering them in their inactive state in the cytosol.

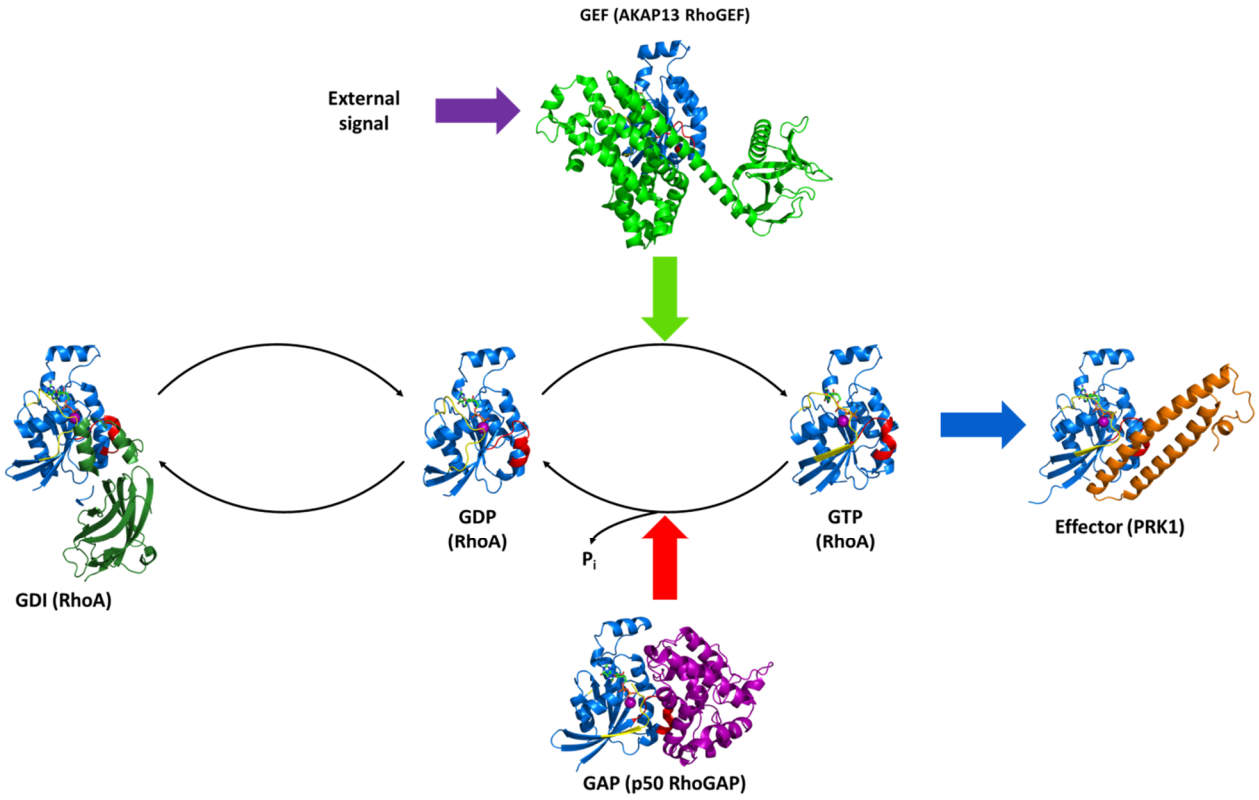


Figure 1.1 G proteins act as molecular switches in signalling pathways. Schematic of the molecular switch mechanism of G proteins. G proteins bind GDP in their inactive state and GTP in their active state. GEFs catalyse nucleotide exchange of GDP for GTP. In their active state G proteins can interact with downstream effectors and hence activate cellular signalling pathways. A GAP interacts with the active G protein and stimulates the rate of hydrolysis of GTP to GDP. P_i is subsequently released and the G proteins revert to their inactive state. A GDI may also negatively regulate some G proteins by removing them from the membrane when they are in their GDP-bound state. GEF (Guanine nucleotide exchange factor), GAP (GTPase activating protein), GDI (Guanine nucleotide dissociation inhibitor). Representative structures used: RhoA·GDP (PDB:1FTN), RhoA·GTPγS (PDB:1A2B), RhoA/AKAP13 RhoGEF (PDB:4D0N, RhoA/p50 RhoGAP (PDB:1TX4), RhoA/PRK1 (PDB:1CXZ), RhoA/RhoGDI-1 (PDB:5FR1).

1.1.1 The G domain

The G domain comprises 5 α-helices and 6 β-sheets which are linked by hydrophobic loops and are arranged in a Rossmann fold with a hydrophobic core (Bourne et *al.*, 1991) (Figure 1.2A). The G domain is conserved in thousands of proteins across all five kingdoms of life (reviewed in Wittinghofer & Vetter, 2011). Five conserved sequence motifs (Table 1.1), termed G1-G5, are part of the core G domain and are important for nucleotide binding and hydrolysis. Figure 1.2B summarises the key contacts between G1-G5 and the nucleotide

(using the residues of H-Ras as the example) while Figure 1.3A illustrates the conserved motifs on the structure of H-Ras.

G1-G4 are strongly conserved within the Ras superfamily while G5 shows weak conservation and in some cases it may be completely absent. G1-G3 are also weakly conserved in proteins which bind to ATP *in lieu* of GTP as they are mostly involved in contacts with the ribose sugar and the phosphate groups of the nucleotide, and G1 and G3 resemble Walker motifs found in other nucleotide-binding proteins (Walker *et al.*, 1982). G4 and G5 instead contact the guanine ring thereby conferring nucleotide selectivity to the G domain. The switch regions are strongly conserved within each of the five families of the Ras superfamily but less conserved between separate families.

Motif	Consensus sequence
G1/P-loop	GXXXXGKS/T
G2	FVDEYDPTIEDSY
G3	DXXGζ
G4	ϕϕϕϕ(N/T)(K/Q)XD
G5	E(A/C/S/T)SA(K/L)

Table 1.1 The G1-G5 conserved sequence motifs. The sequence of each motif is listed in the table. X for any residue, ϕ for hydrophobic residue and ζ for hydrophilic residue.

G1 (also known as P-loop or phosphate-binding loop) contains Gly(15) (H-Ras numbering) and Lys(16) residues which hydrogen bond to the β- and γ-phosphates. In addition the Ser(17)/Thr residue contributes to the coordination of the Mg^{2+} ion which offsets the negative charges of the nucleotide phosphates to allow binding.

G2/Switch 1 is part of β2 and the loop before it and contains two important residues; a Val(29) which hydrogen bonds to the ribose and an invariant Thr(35) which hydrogen bonds to the deprotonated hydroxyl group of the γ-phosphate (this hydrogen bond is broken during GTP hydrolysis). The side-chain hydroxyl group also stabilises the Mg^{2+} ion.

G3 is in part of β3 and the loop that follows and contains a critical Asp residue that hydrogen bonds to a water molecule which is in turn coordinated by the Mg^{2+} ion. G3 also overlaps with switch 2; a glycine residue forms a hydrogen-bond to the GTP γ-phosphate

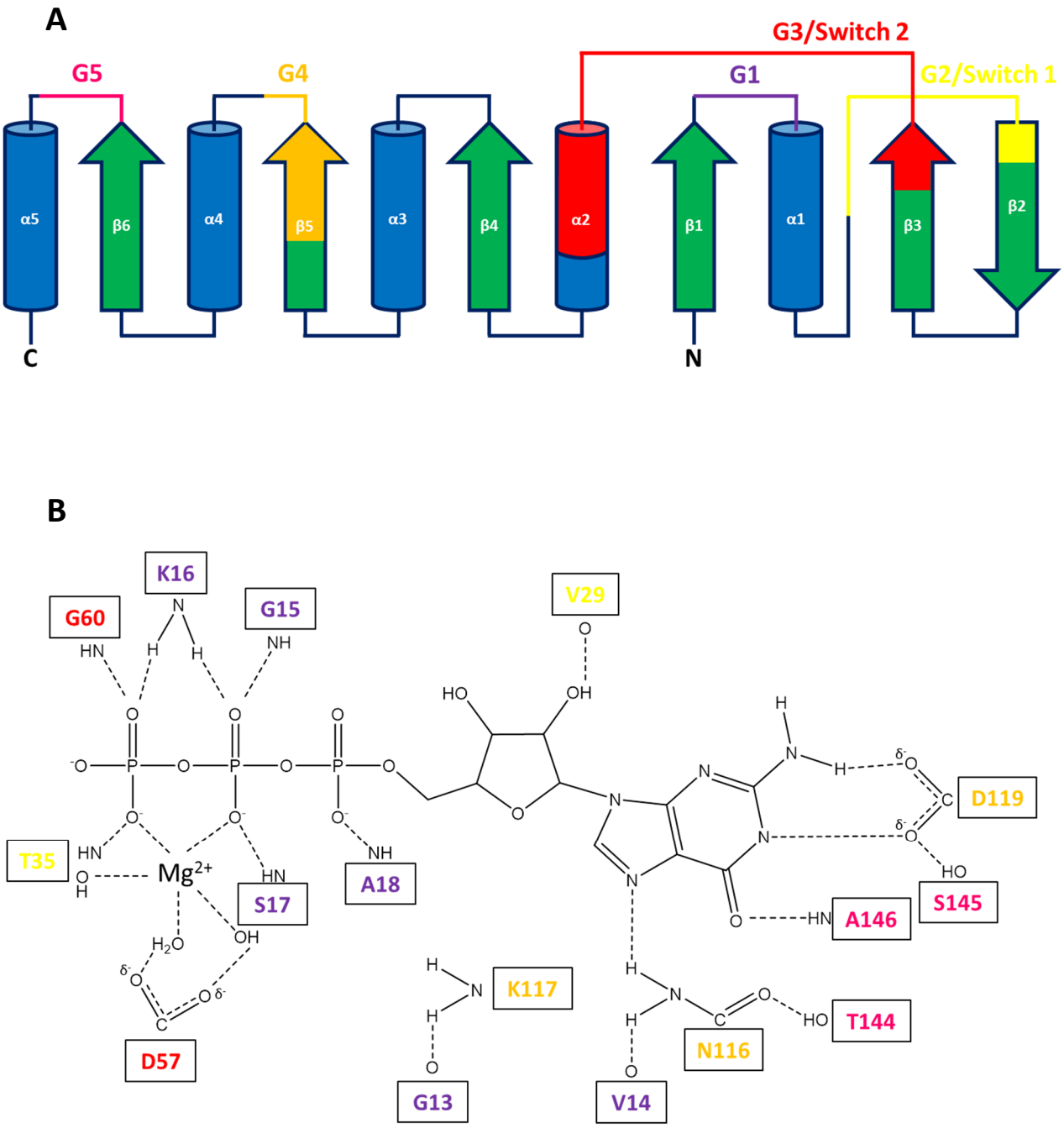


Figure 1.2 The G domain. (**A**) Topology of the G domain. The α-helices are represented by blue cylinders and the β-sheets are represented by green arrows. Dark blue loops connect the α-helices and the β-sheets from N- to C-terminus. Conserved motifs are highlighted; G1 in purple, G2/Switch 1 in yellow, G3/Switch 2 in red, G4 in orange and G5 in magenta. (**B**) Interactions between the G domain and the GTP nucleotide. Key residues involved in the contact were derived from the structure of H-Ras (PDB:5P21). The contacts are described in Bourne *et al.*, 1991 and Wittinghofer & Vetter, 2011. The residues are colour-coded in the same way as the conserved motifs in (**A**).

and is therefore, unsurprisingly, invariant. The hydrophilic residue (Gln61 in H-Ras) is involved in GTP hydrolysis. The direct contact of G2/switch 1 and G3/switch 2 with the γ-phosphates confers the ability to sense the nucleotide state and different conformations are adopted as a result, turning the GTPase into a molecular switch.

G4 is in β5 and in part of the loop before α4 and comprises a stretch of four hydrophobic residues followed by an Asn/Thr and other polar or charged residues and has a key role in nucleotide specificity. The Asn/Thr has a role in stabilising Val14 of G1 while Lys117 also stabilises a G1 residue, Gly13. The Asp hydrogen bonds to the guanine ring, providing nucleotide specificity.

G5 is found in the loop after β6. It contains an Ala which forms an important hydrogen bond to the carbonyl of the guanine ring, an interaction which serves a dual role; specificity to guanine nucleotides and steric exclusion of adenine nucleotides.

1.1.2 The switch regions

The switch regions have an intrinsic flexibility which allows their structure and dynamics to be altered depending on the nucleotide which is bound by the GTPase. Figure 1.3B uses RhoA·GDP and RhoA·GTPγS structures to highlight this. It is this feature which makes effectors bind to GTPases only in a GTP-dependent manner and therefore defines the GTPases as molecular switches. GTP hydrolysis causes the Thr in G2/Switch 1 (and the adjacent Tyr at -3) to lose contact with the nucleotide and the switch thus rotates away from the nucleotide, inducing a conformational change. The same phenomenon occurs with the Gly in G3/Switch 2. The Thr and Gly contacts with the γ-phosphate are often simplistically described to be in a 'tense state' which is lost when GTP is hydrolysed; the so-called 'loaded spring model' (reviewed in Wittinghofer & Vetter, 2011). This is the basis for how effectors can interact with GTPases in a nucleotide-dependent manner and thus transmit signals downstream in signalling pathways.

The inherent flexibility of the switch regions on the μs-ms time scale has been described by NMR relaxation studies (Ito *et al.*, 1997). This flexibility of the switches makes it difficult to

structurally characterise the conformational changes they undergo by NMR or crystallography. For example, peaks for switch II of the small GTPase RheB cannot be seen in NMR spectra as a result of the extensive line broadening which occurs (Schwarten *et al.*, 2007). This is indicative of the switches adopting multiple conformations (Karassek *et al.*, 2010), a property often described as 'regional polysterism' of the GTPases (Ito *et al.*, 1997). This is further supported by ^{31}P NMR which has used the phosphates of the nucleotide to detect switch conformations that have not been detected in crystal structures (Geyer *et al.*, 1996). Importantly, effectors can stabilise a single conformation, e.g. the switches of H-Ras are stabilised by binding of its effector, Raf (Ito *et al.*, 1997), suggesting that effectors and regulators may each recognise, and in turn stabilise, distinct switch conformations.

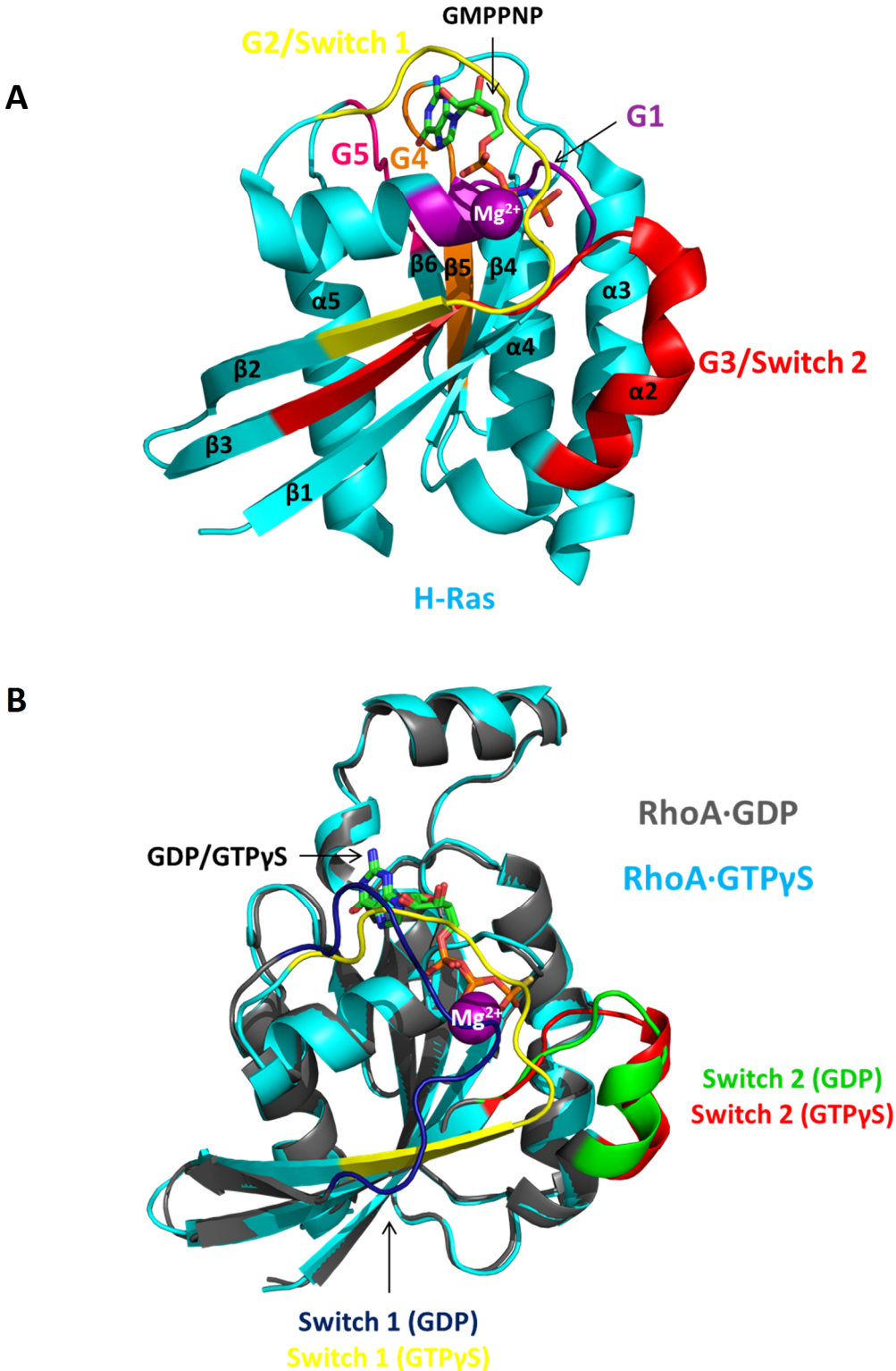


Figure 1.3 Structural features of the G domain. (**A**) The crystal structure of H-Ras·GMPPNP (4L9W) is shown in cyan. Conserved regions are highlighted as follows - G1 in purple, G2/switch 1 in yellow, G3/switch 2 in red, G4 in orange and G5 in magenta. The Mg^{2+} ion appears as a purple sphere. The GMPPNP nucleotide is also labelled. (**B**) Structural comparison of inactive and active conformations of a GTPase using RhoA·GDP (PDB:1FTN) and RhoA·GTPγS (PDB:1A2B) as examples. RhoA·GDP is coloured grey and RhoA·GTPγS is coloured cyan. Switch 1 is coloured navy (GDP) or yellow (GTPγS). Switch 2 is coloured green (GDP) or red (GTPγS).

1.1.3 GEFs

GTPases hold on to GDP very tightly and therefore require the help of guanine nucleotide exchange factors (GEFs) to catalyse exchange of GDP for GTP and activate the GTPase (reviewed in Bos *et al.*, 2007; Cherfils & Zeghouf, 2013). GEFs act by reducing the affinity of the GTPase for GDP in a two-step mechanism which firstly involves binding to the GTPase, and secondly GTP/GDP exchange. The GEF contacts the GDP-bound GTPase at first in a low affinity complex which becomes a high affinity one when GDP dissociates and the GTPase is nucleotide-free. GTP, at ten times higher the cellular concentration of GDP, then binds to the GTPase and causes the GEF to dissociate.

GEFs of different families are structurally unrelated but they all adopt the same mechanism of nucleotide exchange. In particular, all GEFs interact with Switch 1, luring it away from the nucleotide and they also interact with Switch 2, stabilising in this way the nucleotide-free GTPase. Many GEFs also make use of a conserved Asp/Glu residue which is thought to electrostatically repel the nucleotide away from the GTPase, while several others use the conserved Lys in G1/P-loop to form a salt bridge with an Asp/Glu residue of either the GEF or Switch 2 of the GTPase itself, stabilising the nucleotide-free GTPase (reviewed in Cherfils & Zeghouf, 2013).

Rho family GTPases are activated by two major GEF families. The first family comprises Dbl homology (DH)-containing GEFs. In addition to the DH domain, these proteins always contain a pleckstrin homology (PH) domain. Worthylake *et al.* (2000) solved the structure of the first DH/PH GEF, that of Tiam1 in complex with Rac1 (Figure 1.4A). The DH domain is a helical bundle with a 'long chair' shape (Worthylake *et al.*, 2000; reviewed in Cherfils & Zeghouf, 2013). The PH domain is connected to the DH domain by a hinge region and as a result can move around. This may explain why it can promote, inhibit or exist independently of the nucleotide exchange, although *in vitro* the DH domain is sufficient for catalysis (Liu *et al.*, 1998). In the case of Tiam1, the PH domain is thought to push the DH domain into a particular orientation that favours its interaction with Rac1 (Baumeister *et al.*, 2006). Regardless, key structural features are conserved between DH/PH GEFs, as indicated by the

direct comparison of Rac1/Tiam1 and the recently solved structure of RhoA/AKAP13 (Abdul Azeez *et al.*, 2014) (Figure 1.4B).

The second major Rho GEF family are the DOCK proteins which have a Dock Homology Region 1 (DHR1) domain which recruits them to membranes and a DHR2 domain which confers Rac1- and Cdc42-subfamily-specific GEF activity (reviewed in Côté *et al.*, 2007). DOCK GEFs characteristically utilise a hydrophobic valine to push the Mg^{2+} ion out of the nucleotide binding site, thereby catalysing nucleotide release.

Differences in some residues may explain the specificity of GEFs for some GTPases. For example, Vav proteins (DH/PH GEFs) specifically interact with conserved Asp and Glu residues of RhoA and Rac1 (in β2 and β3) to recognise the GTPase and subsequently interact with a conserved lysine (in β5) to catalyse the nucleotide exchange (Movilla *et al.*, 2001; reviewed in Snyder *et al.*, 2002).

GEFs are themselves regulated in several ways, e.g. phosphorylation of the Vav proteins by Src kinases augments their catalytic activity towards RhoA and Rac1 (Crespo *et al.*, 1997). Autoinhibitory, intramolecular interactions between DH domains and SH3 domains have also been reported in the Rac1 GEF, Asef (Murayama *et al.*, 2007). This regulation of GEFs is of utmost importance, as GEFs should only be activated when the appropriate upstream signal is available.

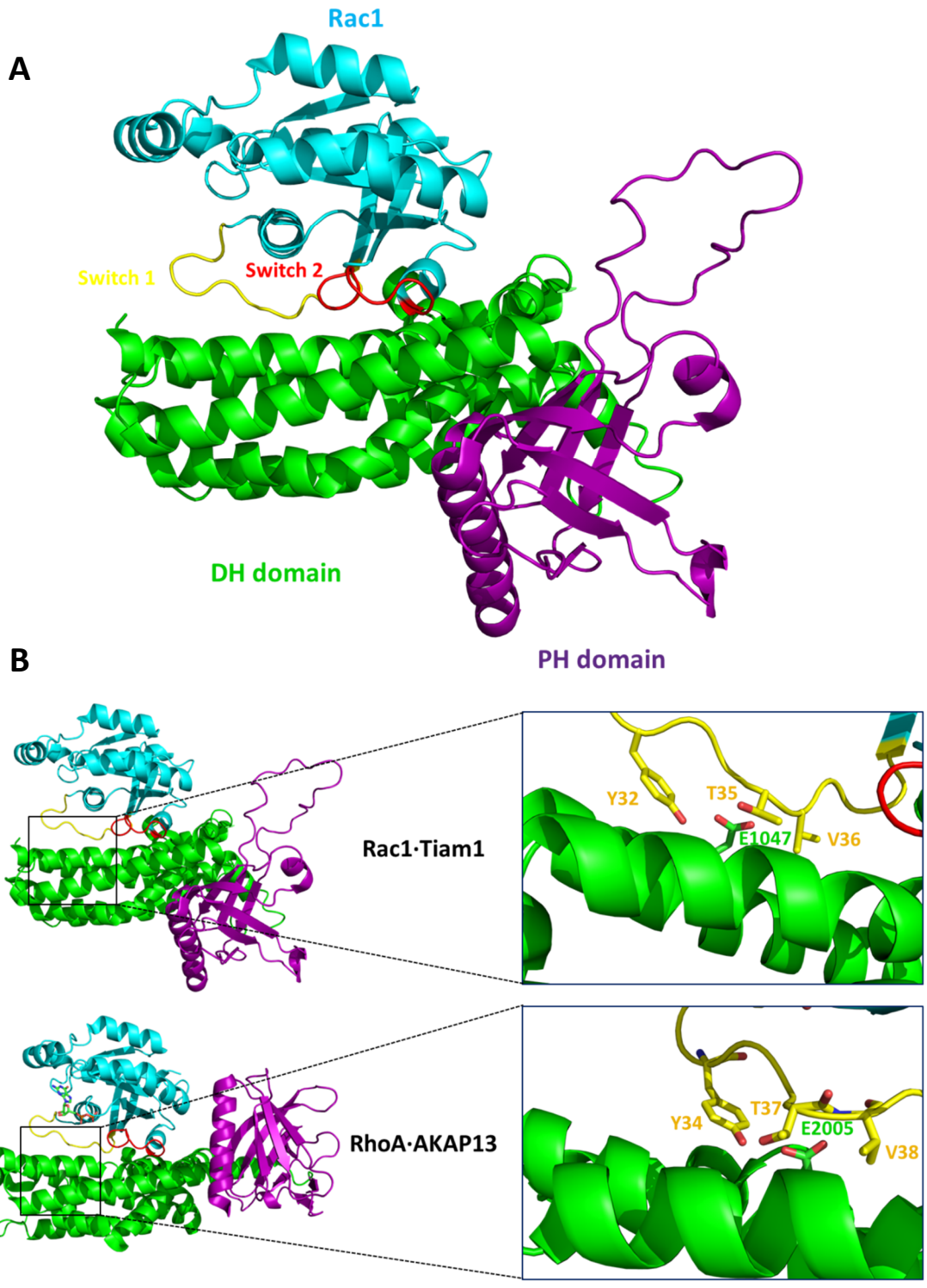


Figure 1.4 Structure and function of GEFs. (**A**) The structure of the Tiam1 (PDB:1FOE) is shown in complex with Rac1. Rac1 is in cyan, switch 1 in yellow, switch 2 in red, the DH domain of the GEF in green, the PH domain of the GEF in purple. (**B**) Structural conservation between the GEFs to allow nucleotide exchange. The structures of Rac1/Tiam1 and RhoA/AKAP13 (PDB:4D0N) are shown here as examples. The conserved catalytic Glu of the RhoGEF is shown to form equivalent interactions with conserved switch 1 residues.

1.1.4 GAPs

GTPase Activating Proteins (GAPs) are also regulators of GTPases. In contrast to heterotrimeric G proteins which hydrolyse GTP quickly, small GTPases exhibit very slow intrinsic GTP hydrolysis rates which are only of 10^{-4}-10^{-5} s^{-1} in magnitude. GAPs stimulate this intrinsic hydrolysis by 10^5 (Lancaster *et al.*, 1994; Lamarche & Hall, 1994), therefore accelerating the return to the inactive, GDP-bound state of the GTPase, which swiftly switches off downstream signalling pathways. While the GAPs of the different families show poor structural conservation (reviewed in Cherfils & Zeghouf, 2013), many GAPs function by inserting an arginine 'finger' into the active site of the GTPases. This has a number of roles. Firstly, the arginine side-chain positions and coordinates the Gln in Switch 2, which in turn coordinates and activates a water molecule for nucleophilic attack, and secondly it offsets the build-up of negative charge from the β- and γ-phosphates, stabilising the transition state and leading to catalysis. This arginine 'finger' mimics the intramolecular arginine which catalyses hydrolysis in heterotrimeric G proteins, making them have much higher intrinsic hydrolysis rates than small GTPases. Rittinger *et al.* (1997) provided the first structure of a RhoGAP, that of the p50 RhoGAP (Figure 1.5A). Figure 1.5B highlights the key components of GAP-catalysed hydrolysis. GAPs can provide completely different hydrolysis rate accelerations even within one family of GTPases, as indicated by a study on RhoGAPs (Amin *et al.*, 2016). Furthermore, specificity of GAPs for particular GTPases may be dependent on other protein-protein interaction domains which are part of RhoGAP domain-containing proteins. GAPs are regulated in a number of ways including intramolecular, autoinhibitory interactions in the case of the RhoGAP β2-chimerin or by membrane curvature in the case of the RhoGAP GRAF or ArfGAP1 (reviewed in Cherfils & Zeghouf, 2013). GAPs and GEFs may also form regulatory complexes to act on distinct GTPases, which may be another way in which the activity of both of these regulators is regulated (reviewed in Hodge & Ridley, 2016).

Without GAPs, the intrinsic rate of hydrolysis is usually faster than the intrinsic rate of nucleotide exchange. This results in a pool of inactive, GDP-bound GTPases which are only

activated in response to upstream signals. Recently however, differences in intrinsic GTP hydrolysis and intrinsic nucleotide exchange rates have been described between members of the Rho family (Jaiswal *et al.*, 2012). This is an example of a shifted paradigm, in which previously thought "canonical" GTPases have features which set them apart from the well-studied members of the family. For example, within the Rho family, RhoD, RhoF, Wrch1 and Wrch2 are considered fast cycling GTPases with high rates of intrinsic nucleotide exchange and low rate of intrinsic nucleotide hydrolysis. Rnd1-3, RhoH and the RhoBTBs on the other hand cannot hydrolyse GTP at all (Jaiswal *et al.*, 2012; reviewed in Aspenström *et al.*, 2007). These 'atypical' GTPases are therefore subject to different layers of regulatory mechanisms, e.g. at the level of expression. These properties have been linked to variations of residues 12, 59 and 61 (Ras numbering) but the residues adjacent to these are conserved making it difficult to define rules under which all of these GTPases fit.

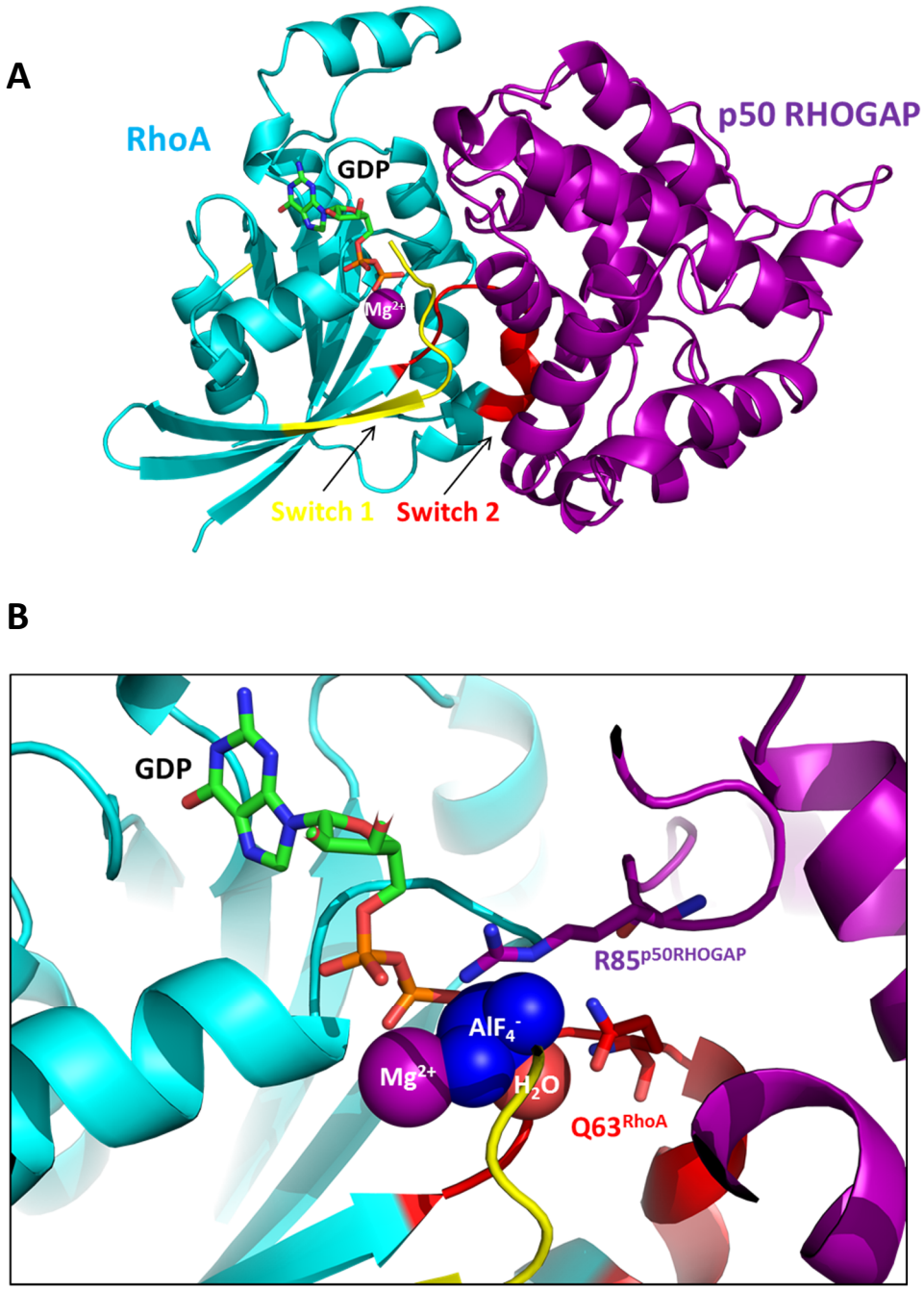


Figure 1.5 Structure and function of GAPs. (**A**) Structure of RhoA in complex with p50 RhoGAP (PDB:1TX4). RhoA is coloured in cyan, switch 1 in yellow, switch 2 in red and p50 RhoGAP in purple. The Mg^{2+} ion appears as a purple sphere. The GMPPNP nucleotide is also labelled. (**B**) Zoomed view of the catalytic site. The 'arginine finger' Arg85 side-chain inserts into the active site and stabilises negative charges during the transition state. Gln63 coordinates and activates the water molecule (in brick red) for nucleophilic attack. The AlF_4^- used to lock the GTPase in the transition state conformation is also highlighted in blue.

1.1.5 GDIs

Guanine nucleotide dissociation inhibitors (GDIs) are additional regulators of the G protein cycle which were first discovered on the basis of their ability to inhibit dissociation of GDP from the GTPase (Fukumoto *et al.*, 1990; Sasaki *et al.*, 1990; Araki *et al.*, 1990). GDIs specific to the Rho and Rab GTPases have been discovered with a total of 3 isoforms per GTPase family. They recognise and bind to the prenyl group of lipid-modified GTPases in their GDP-bound form. They then sequester GTPases in the cytosol which abrogates their ability to interact with downstream effectors.

RhoGDIs have two domains, an N-terminal α-helical bundle that contacts the switches of the Rho GTPase and a C-terminal β-sandwich domain that binds the lipid moiety of the GTPase. Figure 1.6 shows the structure of Rac1 in complex with RhoGDI-1 (Grizot *et al.*, 2001). It is speculated that the interaction of the GDI with the GTPase is stepwise; first the switch regions are recognised by the N-terminal domain, which becomes more structured upon binding to the GTPase (Gosser *et al.*, 1997); this then facilitates binding of the lipid group and thus the subsequent extraction from the membrane (Nomanbhoy *et al.*, 1999). Binding of RhoGDIs to the GTP-bound form of Rho GTPases has been reported (Forget *et al.*, 2002), however the significance of this is not well understood. It is also speculated that RhoGDIs may maintain a pool of Rho·GTP in the cytosol (reviewed in Cherfils & Zeghouf, 2013). Structural differences between the 3 RhoGDI isoforms may also define their differential preference for Rho GTPases. For example, the unique N-terminal amphipathic helix of RhoGDI-3 has been reported to be indispensable for binding to RhoG (Brunet *et al.*, 2002).

Similar to RhoGDIs, RabGDIs also have two distinct domains to bind the GTPase and the geranylgeranyl group. However these domains are structurally dissimilar to those of RhoGDIs (Pylypenko *et al.*, 2006). This may explain the ability of RabGDIs to only bind to Rab·GDP (Wu *et al.*, 2007), in contrast to RhoGDIs which can bind to Rho GTPases loaded with either GDP or GTP. While classical GDIs have not been discovered for the Ras and Arf families, phosphodiesterase δ (PDEδ) can bind to farnesyl groups of GTPases with a β-sandwich domain that resembles that of RhoGDIs (Hanzal-Bayer *et al.*, 2002). PDEδ however lacks a GTPase-interacting motif.

Binding of the GDIs to GTPases leads to the steric occlusion of GEFs and GAPs which ensures that nucleotide exchange and hydrolysis are prevented. This ensures the fidelity of the molecular switch mechanism, which needs to be tightly controlled. GDIs are regulated in several ways including phosphorylation, e.g. protein kinase C phosphorylates RhoA to reduce its affinity for RhoGDI, while protein kinases A and G phosphorylate RhoA to stabilise this interaction (Dovas *et al.*, 2010; Rolli-Derkinderen *et al.*, 2005; reviewed in Cherfils & Zeghouf, 2013). The effect of these phosphorylation events cannot be rationalised based on structures of Rho GTPases in complex with GDIs; phosphorylation may be inducing allosteric effects in the GTPase. SUMOylation can also regulate GDIs as indicated by SUMOylated RhoGDI binding more tightly to Rac1 (Yu *et al.*, 2012).

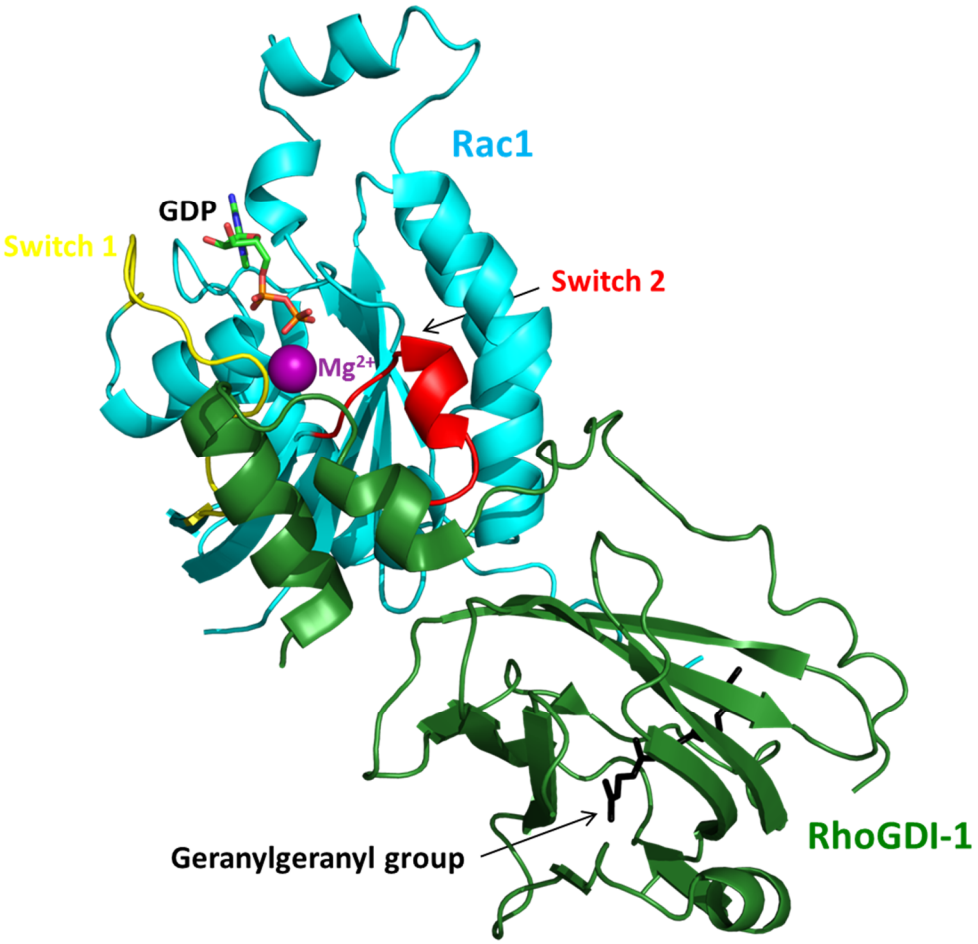


Figure 1.6 Structure of Rac1 in complex with RhoGDI-1. Rac1 is shown in cyan and RhoGDI-1 in dark green. Switch 1 (yellow), switch 2 (red), GDP, Mg^{2+} (purple sphere) and the geranylgeranyl group of Rac1 (black) are highlighted (PDB:1HH4).

1.1.6 Effectors

Effectors are molecules which interact with GTPases in a GTP-dependent manner. Binding of the effector to the GTPase often leads to the activation of the effector which can then mediate cellular signals in downstream pathways. It is possible for more than one effector to interact with each GTPase, effectively implicating GTPases in a plethora of signalling pathways and cellular effects. Each of the 5 families of GTPases within the Ras superfamily can be assigned to a particular cellular effect that is mediated through its effectors; Ras proteins control cell growth and differentiation, Rho proteins control cytoskeletal reorganisation, Rab and Arf proteins are involved in vesicular transport and Ran, the only member of its family, controls nuclear import and export.

Dissecting signalling pathways and understanding the role of GTPase/effector interactions remained a challenging task until the advent of structural studies. The first structure of a GTPase in complex with an effector was Rap1A (a Ras family GTPase) in complex with Raf (Nassar *et al.*, 1995). Since then, ~70 structures of GTPase/effector complexes have been solved, indicating that there are three main groups of structural motifs which effectors contain that bind to the GTPase, and particularly to switch 1, switch 2, the interswitch region (consisting of parts of α1 and β1) and β3. The largest group of effectors is that which interacts with GTPases using a pair of α-helices. The effectors in this group can be sub-divided into six classes based on the orientation of the helices with respect to each other and the GTPase switch regions; these classes comprise four anti-parallel coiled coil classes and two parallel coiled coil classes (reviewed in Mott & Owen, 2015). The second group is that which uses intermolecular β-sheet formation for the interaction, while the third and smallest group is that of effectors using a PH domain to bind to the GTPase. Structures show that PH domains all use different interfaces and secondary structure elements to interact with different GTPases (reviewed in Mott & Owen, 2015), illustrating that GTPase/effector interactions are far from predictable but instead are an enigma.

Despite the large number of known effectors, it is likely that there are many effectors which have not yet been identified. This is highlighted by the less well-studied GTPases for which only a few effectors have been identified. Lack of conserved binding motifs across the

different families makes their identification difficult. Recent studies however have identified some new interactions (Paul *et al.*, 2017) but also highlight the tremendous lack of our understanding of which signalling pathways are governed by which GTPases.

Effectors may be activated directly by the GTPase, e.g. the GTPase relieves effector autoinhibition by inducing a conformational change, or may be activated by being localised to the correct cellular compartment by the GTPase. An example of direct activation by a GTPase is that of the PAK1 kinase which forms inactive homodimers that are dissociated upon activation by Rac1 and Cdc42 (Parrini *et al.*, 2002).

1.1.7 Post-translational modifications

Ras, Rho and Rab GTPases are prenylated at their C-terminus by the mevalonate pathway. Ras family proteins are modified with the addition of a farnesyl group (Schafer *et al.*, 1989; Casey *et al.*, 1989). However, while K-Ras is only farnesylated, H-Ras can also be palmitoylated. Rho family proteins are often modified with a geranylgeranyl group (Katayama *et al.*, 1991), although RhoD, TCL, TC10, Rnd1, Rnd2 and Rnd3 are exclusively farnesylated (Roberts *et al.*, 2008; Foster *et al.*, 1996). Rab proteins are also geranylgeranylated (Horiuchi *et al.*, 1991) and this is often a double geranylgeranylation on two C-terminal cysteines (Seabra *et al.*, 1992). In contrast, Arf proteins are myristoylated at the new N-terminus following the removal of the N-terminal methionine (Lodge *et al.*, 1994). Ran is not modified at all, which allows it to move freely in and out of the nucleus. Figure 1.7A illustrates the different lipids that can be added to GTPases.

Prenylation of Ras, Rho and Rab GTPases is catalysed by either one of two prenyl transferases, farnesyl transferase (FTase) or geranylgeranyl transferase (GGTase) (Figure 1.7B). These recognise a C-terminal motif known as the CAAX box (C=cysteine, A=aliphatic, X=any). The specificity for FTase or GGTase is determined by the X residue, which is the last residue in the protein. The prenyl transferases add the lipid group to the cysteine of the CAAX box. Subsequently, the AAX sequence is cleaved by Ras-converting CAAX endopeptidase 1 (RCE1) which is followed by the methylesterification of the cysteine by

isoprenylcysteine carboxymethyltransferase (ICMT). Palmitoylation is catalysed by palmitoyltransferase enzymes, e.g. H-Ras is palmitoylated by a heterodimeric complex consisting of the ZDHHC9 and GCP16 proteins (Swarthout *et al.*, 2005).

Prenylation has several effects on GTPases. Firstly, it is important for the correct cellular localisation of the GTPase, targeting the GTPase to the right membrane where it can interact with the correct regulators and effectors (Roberts *et al.*, 2008; Taguchi & Misaki, 2011). A study using an inhibitor specific to GGTase demonstrated that this is enough to prevent the activation of RhoA and as a result cause the disruption of the actin cytoskeleton (Khwaja *et al.*, 2006). Secondly, the prenyl group is recognised directly by the GDIs, allowing them to extract GTPases from the cell membranes. Thirdly, it has been shown to have an effect on GEF and GAP activity (Pechlivanis *et al.*, 2007; Moskwa *et al.*, 2005). The affinity of GTPases for their effectors is often determined using soluble proteins; however it is likely that the anchored GTPases bind differently to effectors in the context of a lipid bilayer. Furthermore, while prenylation does not appear to have a structural effect on GTPases such as Ras (Güldenhaupt *et al.*, 2008), it may induce structural changes in other GTPases. For example, NMR studies have shown that the myristoyl group of Arf6 folds back to interact with hydrophobic residues in the N-terminal domain of the protein (Gizachew & Oswald, 2006). The GTPases may also associate with the membrane at different orientations which may dictate how they interact with downstream effectors (reviewed in Prakash & Gorfe, 2017). The significance of prenylation is highlighted by RCE1 and ICMT knockout mice which show abnormal development (reviewed in Wang & Casey, 2016). RhoB remains an interesting case, as it can be farnesylated or geranylgeranylated, states which are linked to RhoB's tumour promoting or tumour suppressing roles, respectively (Du & Prendergast, 1999). The identity of the prenyl group is therefore sufficient, at least in the case of RhoB, to determine the function of the protein (by determining its localisation, i.e. plasma membrane or endosome).

Small GTPases are also subject to other types of post-translational modification, including phosphorylation, SUMOylation and ubiquitylation (reviewed in Hodge & Ridley, 2016). For example, phosphorylation of Tyr64 and Ser71 of Rac1 (both in switch 2) by Src and PKB, respectively, leads to stabilisation of the GDP-bound state of the GTPase with consequential effects on the actin cytoskeleton (Chang *et al.*, 2011; Kwon *et al.*, 2000). In contrast,

phosphorylation of Thr108 leads to nuclear translocation (Tong *et al.*, 2013). Therefore distinct phosphorylations mediate distinct regulatory effects on GTPases.

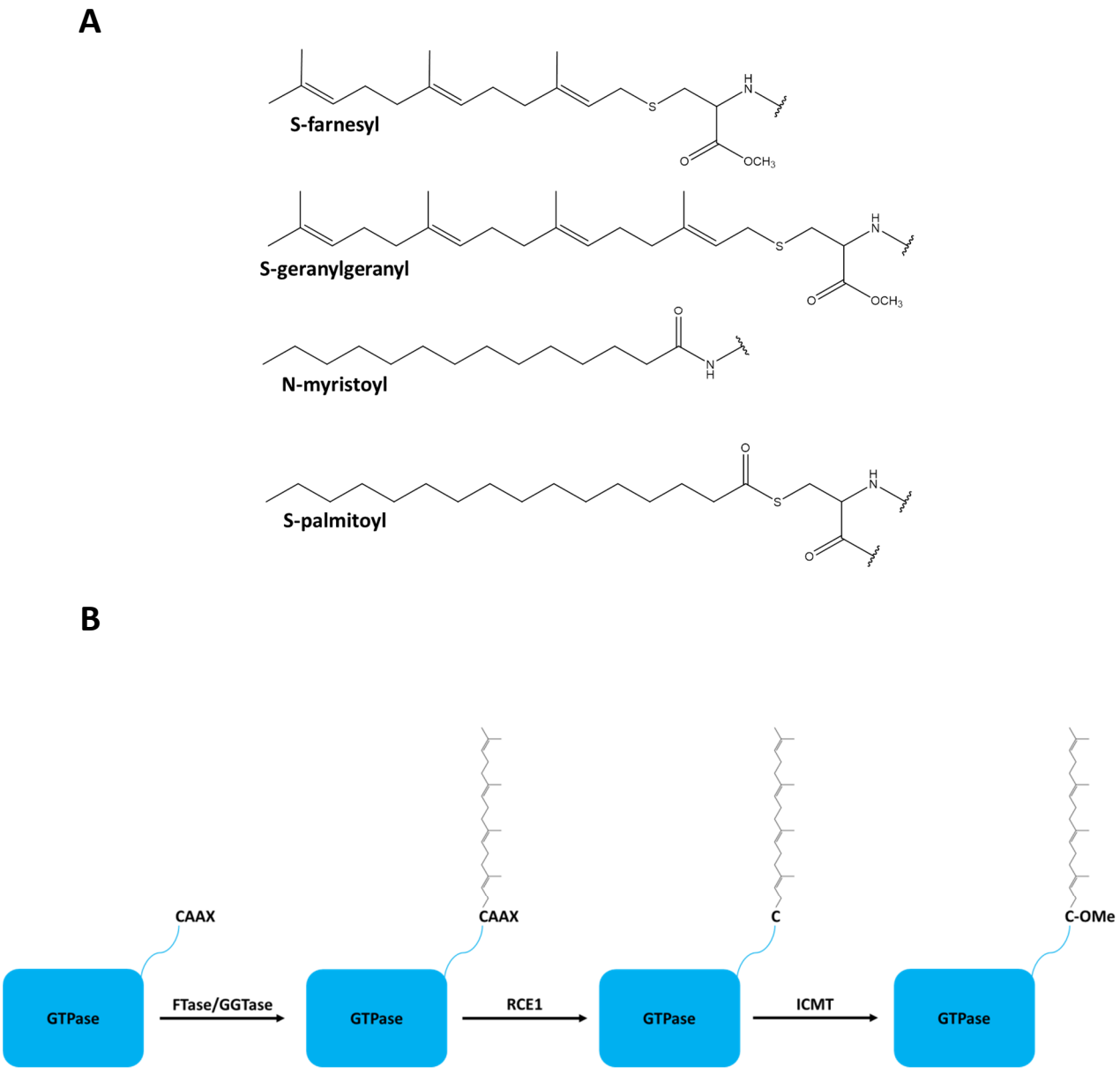


Figure 1.7 Prenylation of GTPases. (**A**) Post-translational modifications of GTPases by lipids. (**B**) Mevalonate pathway for the farnesylation or geranylgeranylation of Ras, Rho and Rab GTPases. Geranylgeranylation is shown as an example. CAAX (C=cysteine, A=aliphatic, X=any) represents the conserved four-amino acid motif at the C-terminus of GTPases. FTase for farnesyltransferase, GGTase for geranylgeranyltransferase, RCE1 for Ras-converting CAAX endopeptidase 1 and ICMT for isoprenylcysteine carboxymethyltransferase.

1.1.8 Structural similarities and differences between the 5 families

The members of the 5 families within the Ras superfamily all share the core G domain. However, there are also distinct structural elements which are unique to each of the Rho, Arf and Ran families (Figure 1.8). Rho GTPases have a 13-residue α-helix located between α4 and β5 of the G domain (Hirshberg *et al.*, 1997). This Rho insert does not contact any regulators or effectors of Rho GTPases but it has been shown to be necessary for cellular transformation (Wu *et al.*, 1998; Zong *et al.*, 2001) and it has also been reported to have a different conformation when a GDI is bound (Abramovitz *et al.*, 2012). Arf GTPases have an N-terminal α-helix which is involved in membrane binding (Goldberg, 1998). When Arf proteins bind to GTP, there is a conformational change involving the antiparallel β-sheet (known as the 'interswitch toggle') which causes the N-terminal helix to be displaced and interact with the membrane (Amor *et al.*, 1994; Pasqualato *et al.*, 2002). On the other hand, the C-terminal extension of Ran serves to define the GDP-bound state with its deletion affecting the switch regions and leading to a constitutively active protein (Nilsson *et al.*, 2002). Specifically, GTP/GDP exchange makes the C-terminal helix no longer able to bind to Ran, allowing it to interact with its effectors (Scheffzek *et al.*, 1995).

While the majority of the GTPases in the Ras superfamily share sequence and structural similarity and are regulated in similar ways, they are involved in a multitude of cellular processes. This diversity, an amalgamation of the effects of subtle structural and regulatory differences, differential binding to effectors and localisation, makes them fascinating to study and highlights their immense complexity.

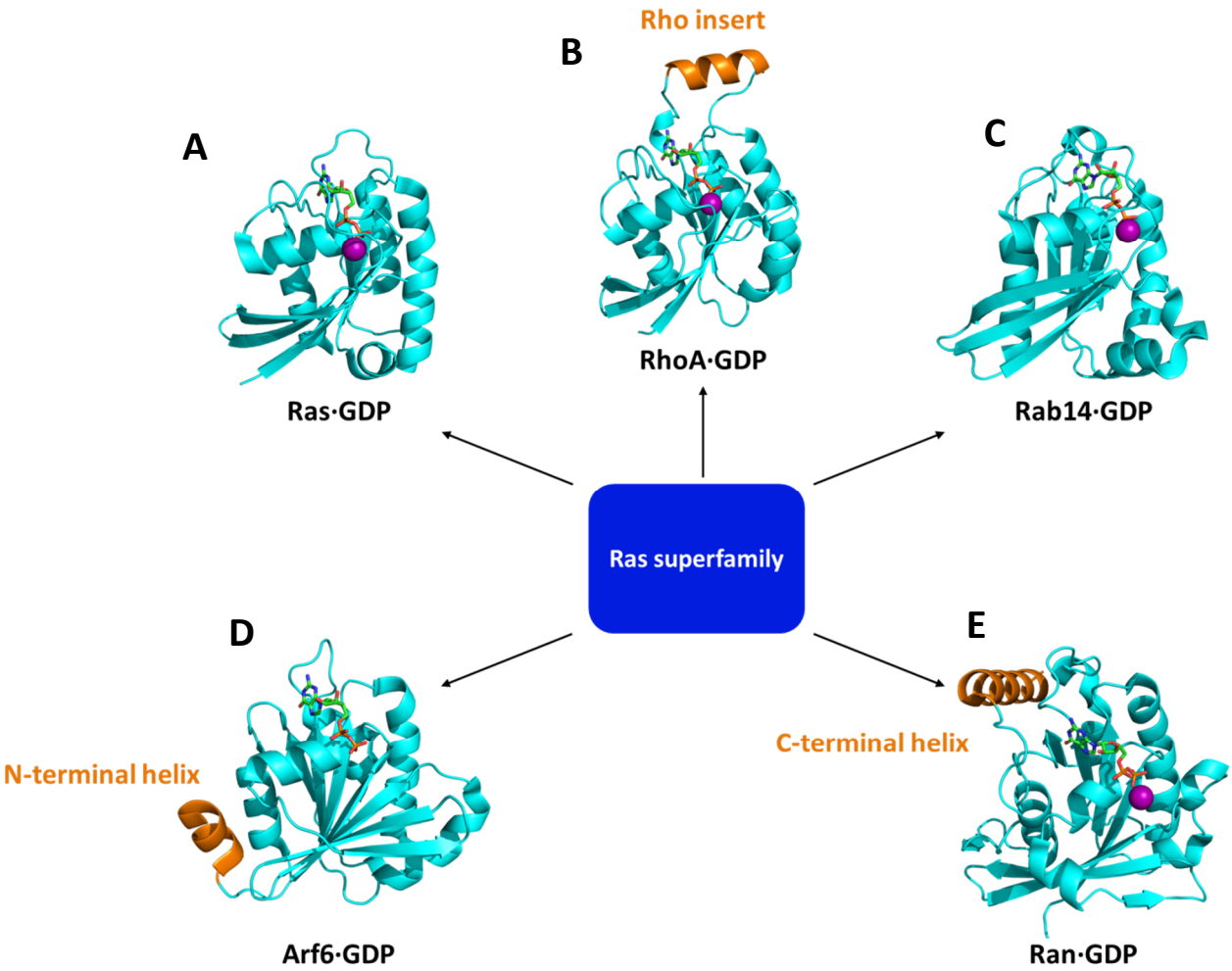


Figure 1.8 The 5 families of the Ras superfamily. Representative structures of a member of each of the families are used to highlight the overall fold as well as any unique structural elements. (**A**) H-Ras·GDP of the Ras family (PDB:4L9S) (**B**) RhoA·GDP of the Rho family (PDB:1FTN) (**C**) Rab14·GDP of the Rab family (PDB:1Z0F) (**D**) Arf6·GDP of the Arf family (PDB:1E0S) (**E**) Ran·GDP (PDB:3GJ0). GTPases are in cyan and any additional structural elements are in orange. GDP and the Mg^{2+} ion (purple sphere) are also highlighted.

1.2 The Rho family

1.2.1 Implications in several signalling pathways and cellular localisation

The Rho family consists of 23 members which can be divided into sub-families based on their evolutionary similarity (Figure 1.9). RhoA, Rac1 and Cdc42 are the three archetypes of the family and their regulators, effectors and physiological roles have been well-characterised. Out of the remaining 20 Rho GTPases, some have not been characterised substantially however it is known that there is a diverse range of PTMs, localisation, effector binding and cellular functions of the Rho GTPases (reviewed in Vega & Ridley, 2007).

Rho GTPases couple changes in the extracellular environment to changes in intracellular signal transduction pathways. Similar to the diverse range of functions across the Ras superfamily, a surprising diversity in function is also observed amongst the Rho family GTPases. This is true even for members within a specific subfamily. For example, the Rho subfamily members RhoA, RhoB and RhoC share 84% sequence similarity. Regulators and some effectors are shared amongst these three GTPases, as are some key biological functions, e.g. all three Rho subfamily GTPases induce stress fibre formation. Despite this, differential roles have been observed for these three proteins, e.g. RhoA is involved in cytokinesis while RhoB regulates endosomal transport, suggesting that they are not redundant. This may be partly due to the different cellular compartments in which they can be found. In particular, RhoA and RhoC localise to the plasma membrane, while RhoB localises to the endosomes. RhoA and RhoC share 91% sequence identity and they are more similar to each other than to RhoB, yet the effects they have on cell morphology are distinct (Vega *et al.*, 2011). This is thought to be a result of them acting through different effectors. Further differences in function are thought to arise from the distinct cellular localisation of these GTPases which is partly dependent on their specific lipid modification. For instance, RhoA, RhoC, Rac1 and Cdc42 are localised to the cell membrane, while RhoD and RhoB are localised to early and late endosomes, respectively. This localisation is often coupled to the corresponding localisation of the effectors, e.g. PRK1 and mDia1 localise to the endosomes where they interact with RhoB (Mellor *et al.*, 1998; Fernandez-Borja *et al.*, 2005).

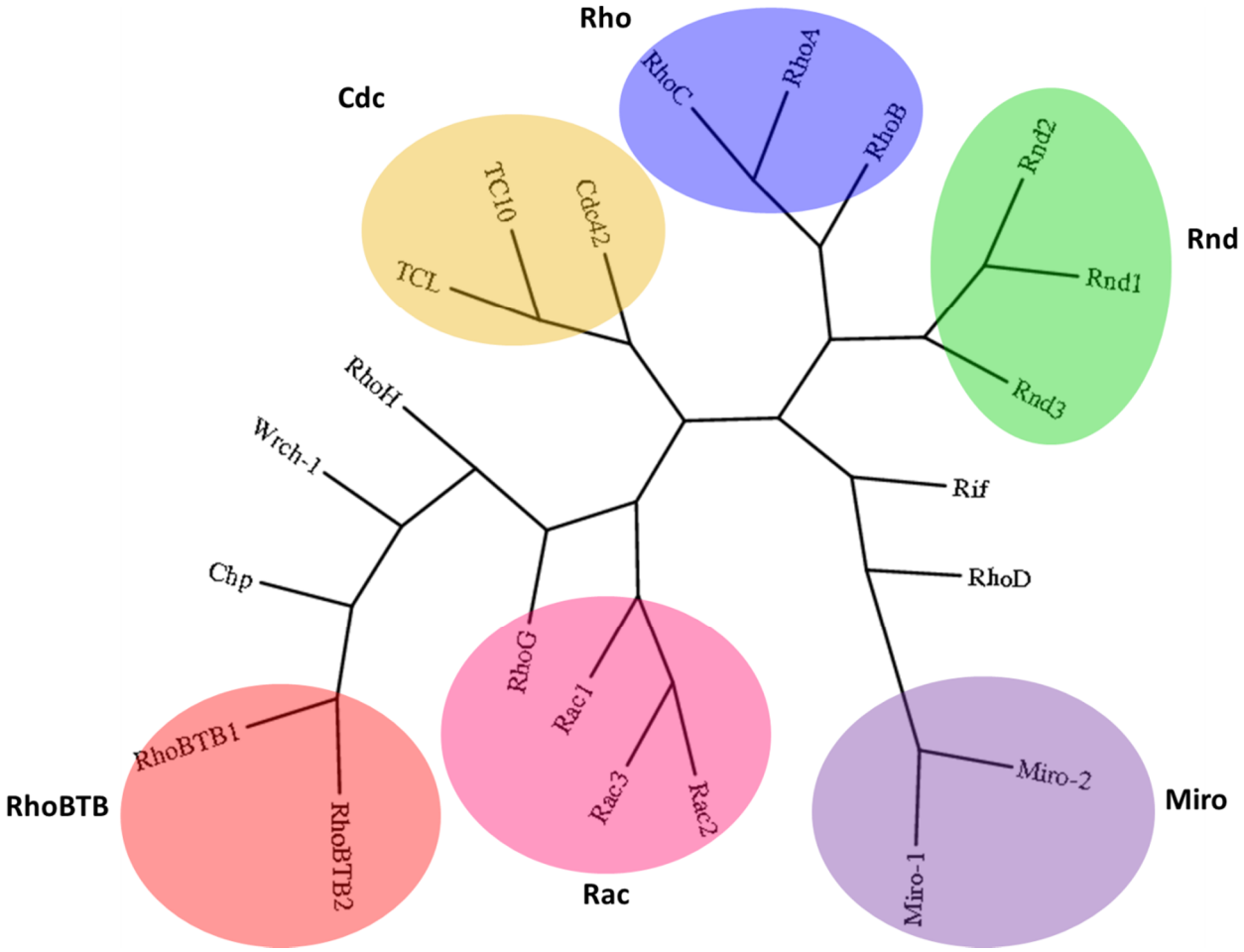


Figure 1.9 Phylogenetic tree of the Rho family GTPases. The different subfamilies are highlighted. The unrooted phylogenetic tree was constructed on phylogeny.fr as described in Dereeper *et al.* (2008).

1.2.2 The three key players

Each of the three key players is responsible for inducing distinctive effects on the cell cytoskeleton. RhoA controls the rearrangement of existing actin and myosin filaments to form stress fibres (Ridley & Hall, 1992), whereas Rac1 and Cdc42 cause the *de novo* polymerisation of actin, leading to formation of cell protrusions known as lamellipodia and filopodia, respectively (Ridley *et al.*, 1992; Nobes & Hall, 1995; Kozma *et al.*, 1995). These effects are governed by effectors which are specific to these GTPases (see section 1.3). The cellular structures regulated by the three archetypal Rho GTPases are essential for cell migration (Figure 1.10) (reviewed in Zegers & Friedl, 2014; Ridley, 2015). The three archetypal Rho GTPases are also involved in the regulation of the cell cycle in yeast (Olson *et al.*, 1995) and control of vesicular trafficking (reviewing in Symons & Rusk, 2003).

1.2.3 Role in disease

The Rho family is implicated in many diseases and notably in cancer. Originally RhoGEFs were discovered as oncogenes, and Rho GTPases were thought to only be deregulated via the RhoGEFs (reviewed in Porter *et al.*, 2016). Recently however, several Rho family mutations have been identified in various cancers, e.g. fast-cycling P29S Rac1 in melanoma (reviewed in Haga & Ridley, 2016; Bustelo, 2018) and while Rho proteins are often described as drivers of metastasis, some proteins like RhoB may have a tumour suppressing role in addition to a tumourigenic one (Ju & Gilkes, 2018). Understanding the structure, function and regulation of Rho GTPases, their regulators and their effectors is key for designing therapeutics to target Rho GTPase-driven pathologies.

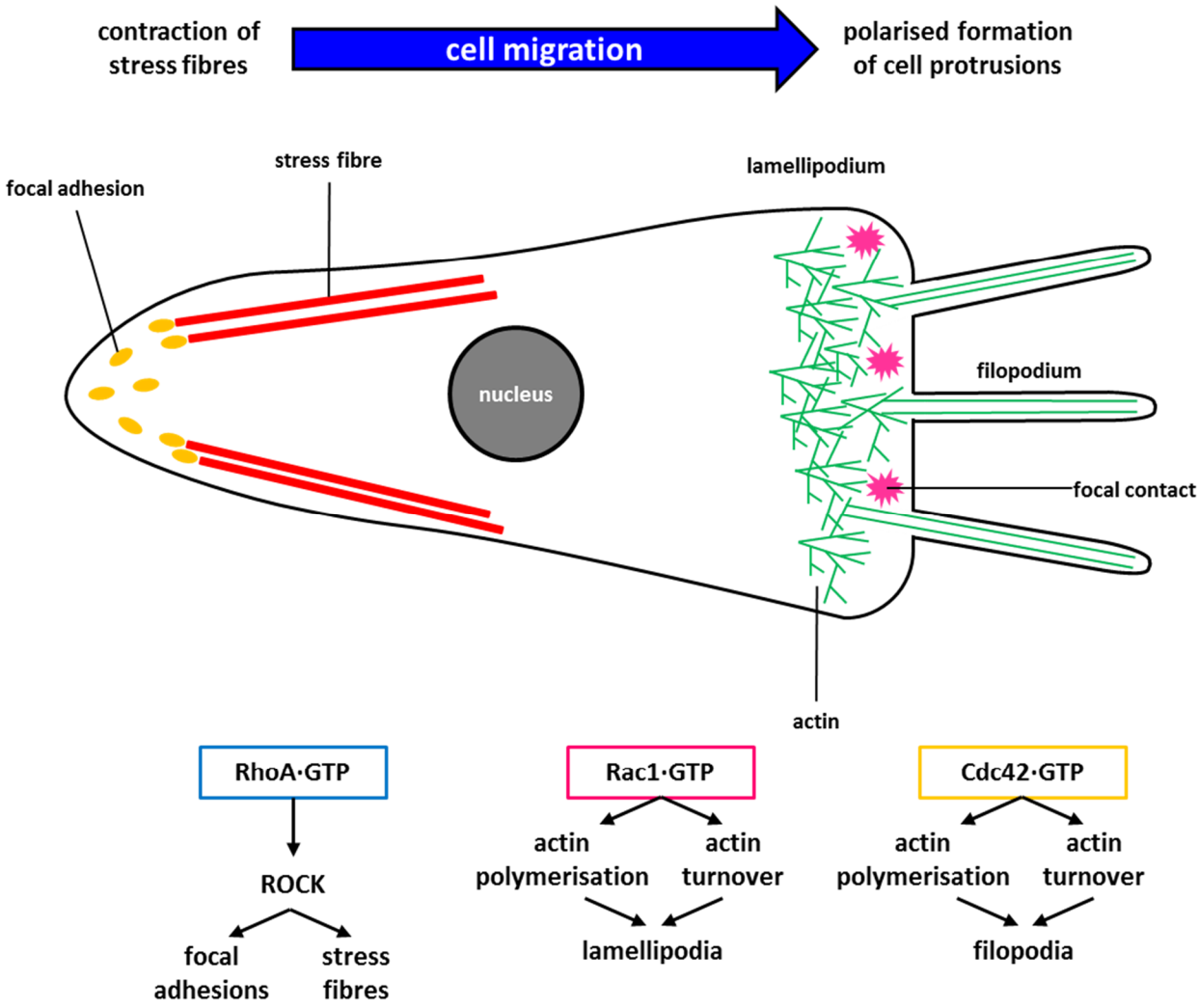


Figure 1.10 Cytoskeletal structures controlled by Rho GTPases. RhoA controls the formation of stress fibres, Rac1 controls the formation of lamellipodia and Cdc42 controls the formation of filopodia. The concerted contraction of stress fibres and polarised formation of cell protrustions allow the cell to migrate.

1.3 Rho family effectors

Rho family GTPases have an abundance of effectors. Over 70 effectors have been identified for each of the archetypal RhoA, Rac1 and Cdc42 GTPases (Iden & Collard, 2008). However our knowledge and understanding of Rho family effectors still remains nebulous, particularly for less well-characterised Rho GTPases. A recent mass spectrometry study identified several novel effectors even for canonical GTPases (Paul *et al.*, 2017), highlighting the missing pieces in the Rho family signalling puzzle. Effectors can be grouped into different categories by means of how they interact with the Rho GTPases (reviewed in Hakoshima *et al.*, 2003; Mott & Owen, 2015). The physiological roles of the key effector families of RhoA, Rac1 and Cdc42 will be briefly described below, as well as any structural features involving their interaction with Rho GTPases.

1.3.1 Rac1 and Cdc42 effectors

1.3.1.1 CRIB motif effectors

Rac1 has been implicated in lamellipodia formation (McCarty *et al.*, 2005) through its interaction with its effector IRSp53 (Miki *et al.*, 2000) and focal contact turnover through its interaction with p21-activated kinase (PAK) (Vidal *et al.*, 2002; Delorme-Walker *et al.*, 2011). In contrast, Cdc42 is involved in filopodia formation, partly by activating Neuronal Wiskott-Aldrich Syndrome protein (N-WASP) which signals to the Arp2/3 actin filament nucleation complex (Rohatgi *et al.*, 2000).

These effectors interact with Rac1 and Cdc42 via a CRIB (Cdc42, Rac interactive binding) motif. The CRIB motifs of the WASP and ACK effectors are unstructured when free in solution and become structured when they interact with Cdc42 and Rac1, forming anti-parallel intermolecular β-sheets (Figure 1.11B). This type of interaction has been seen with Ras binding to its effector Raf, however while in that case the interaction is driven by electrostatic contacts, CRIB motif effectors also make a series of hydrophobic contacts with the GTPase. An additional difference is that these CRIB motifs adopt an extended

conformation, making additional contacts with switch 2 that are absent in the Ras-Raf complex.

1.3.1.2 HR1 effectors

The Transducer of Cdc42-dependent actin assembly protein 1 (TOCA1) and the closely-related proteins CIP4 and FBP17 are effectors of Cdc42 which are involved in actin polymerisation. They interact with active Cdc42 via their HR1 domain which shows sequence and structural homology to the PRK HR1 domains, and they also have an SH3 domain and an F-BAR domain (Ho *et al.*, 2004; Watson *et al.*, 2016).

The HR1 domains of this family interact with Cdc42 with low μM affinities, in contrast to PRK HR1 domains which have nM affinities for GTPases. Differences in the inter-helical loop may explain this weaker binding. The transient protein-protein interactions of the TOCA1 family are thought to be critical for the recruitment of N-WASP in an 'effector handover' that facilitates actin polymerisation (Watson *et al.*, 2016).

1.3.1.3 IQGAP

IQGAPs (IQGAP1 and IQGAP2) interact with Rho GTPases via their GAP-like domain, which cannot stimulate GTP hydrolysis by the GTPase as it lacks the key arginine finger (Kurella *et al.*, 2009). The other domains of IQGAP implicate it in several cellular processes, including actin cytoskeletal reorganisation, calmodulin signalling, E-cadherin-mediated adhesion and a role as a scaffold for components of the MAPK pathway (reviewed in Watanabe *et al.*, 2015).

IQGAP is an effector of both Rac1 and Cdc42 and mutational analysis has revealed that these G proteins use overlapping but not identical interaction interfaces (Owen *et al.*, 2008). Furthermore, Cdc42 was shown to bind more tightly to an extended GAP-like domain than to the isolated GAP domain (Owen *et al.*, 2008). Interestingly, Cdc42, but not Rac1, binds to this additional site in IQGAP resulting in the stabilisation of IQGAP dimers (LeCour Jr. *et al.*, 2016). This highlights key differences on the effects that different GTPases may have on the same binding partner.

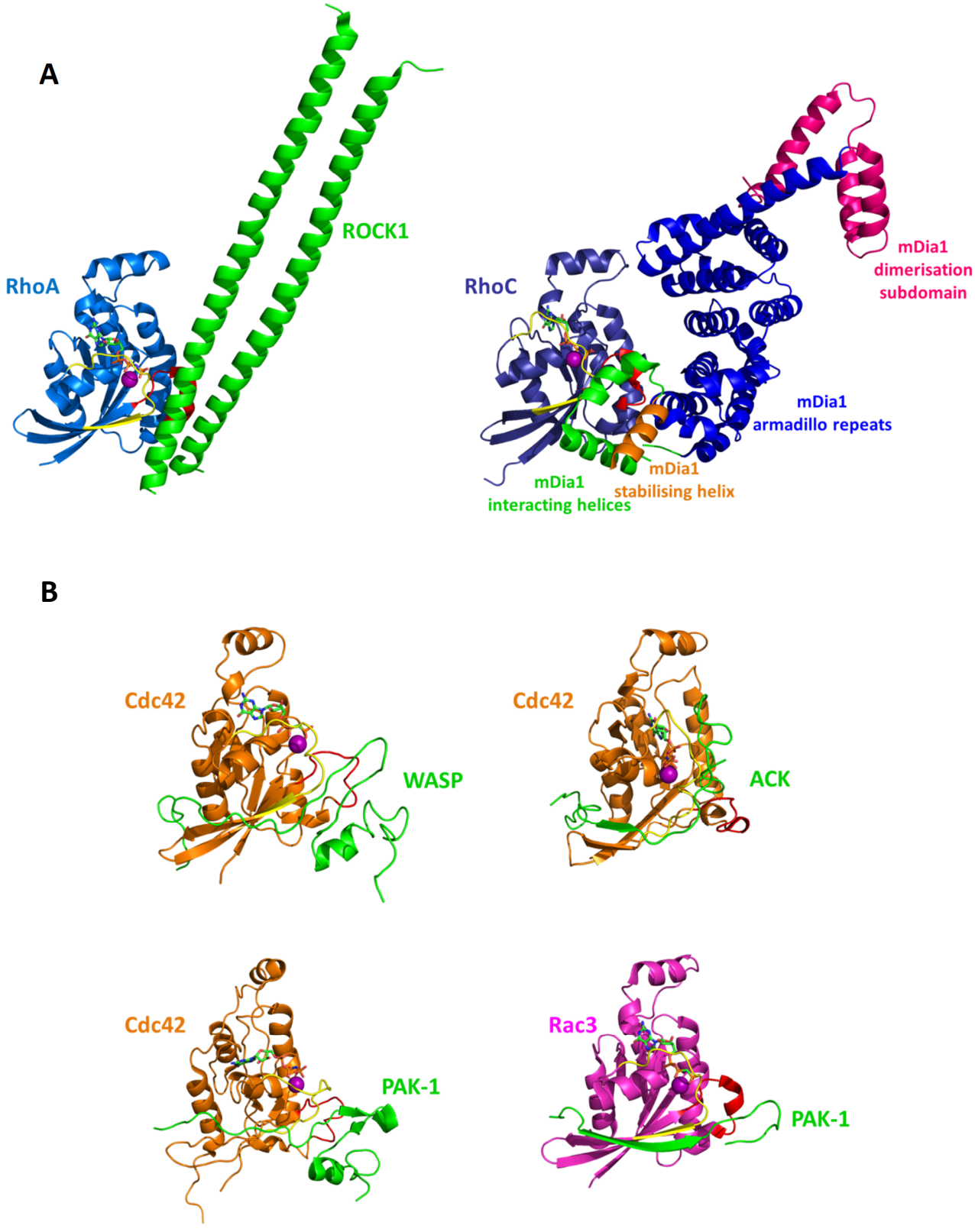


Figure 1.11 Rho GTPase/effector complex structures. The effectors are categorised based on the fold of the effector which interacts with the switch regions of the GTPase. (**A**) α-helical effectors include RhoA/ROCK1 (PDB:1S1C) and RhoC/mDia1 (PDB:1Z2C). (**B**) The CRIB effectors of Rac1 and Cdc42. Shown here are structures of Cdc42/WASP (PDB:1CEE), Cdc42/PAK (PDB:1E0A), Cdc42/ACK (PDB:1CF4), Rac3/PAK1 (PDB:2QME).

1.3.2 RhoA effectors

The key RhoA effector families and the cellular processes they control are outlined in Figure 1.12.

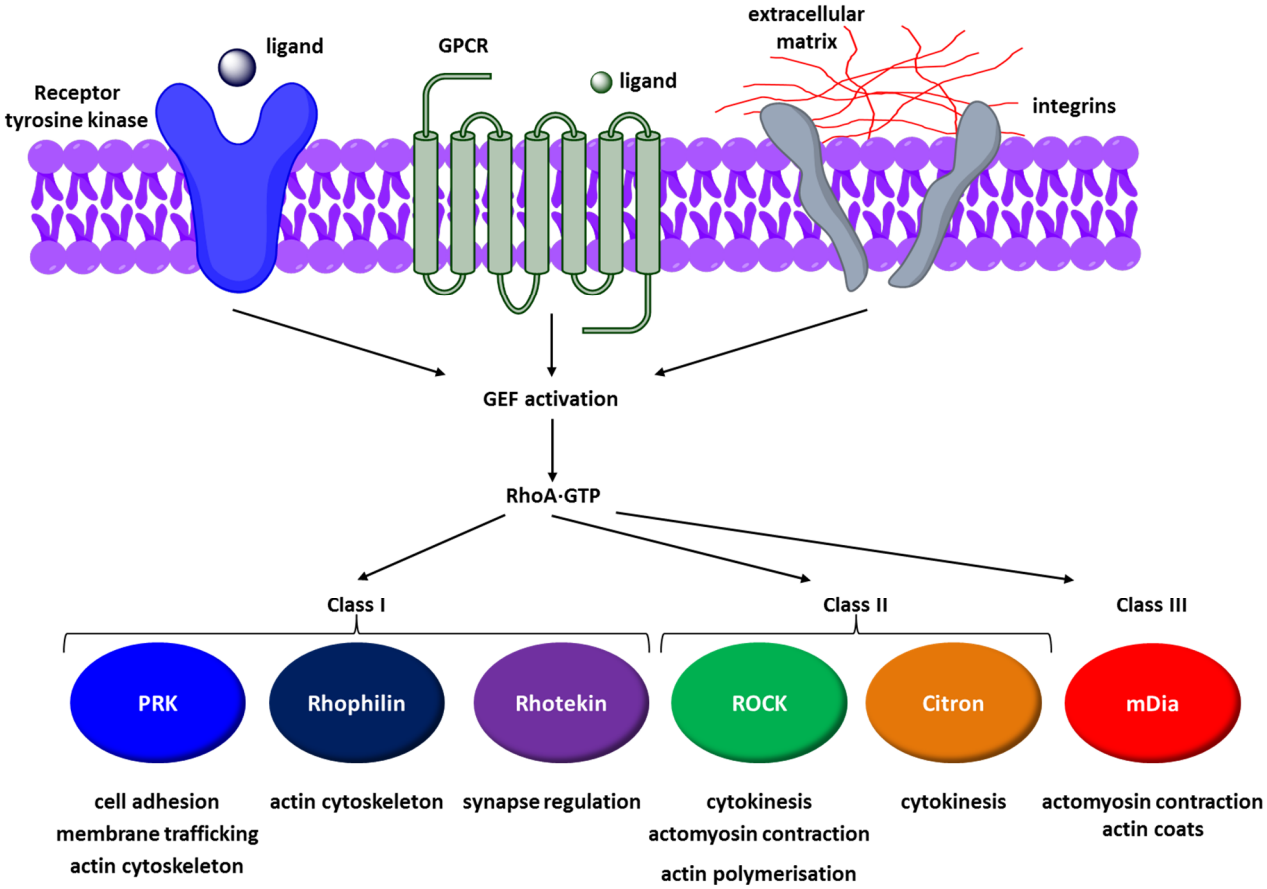


Figure 1.12 Summary of the key effectors of RhoA. The diagram summarises the main types of receptors which activate GEFs, leading to the activation of RhoA. RhoA then binds to the Rho-binding domain of several downstream effectors which can be grouped into different classes based on sequence and structural homology. These effectors control distinct cellular functions.

1.3.2.1 The diaphanous proteins

The diaphanous proteins, or mDia (mammalian homolog of Drosophila diaphanous) proteins are Rho effectors which contain formin-homology domains in addition to a Rho-binding domain (RBD). They are involved in the formation of actomyosin bundles such as stress fibres and the contractile ring (reviewed in Narumiya *et al.*, 2009; Thumkeo *et al.*, 2013). mDia1 also co-localises to endosomes with RhoB, where it promotes the formation of an actin coat around endosomes (Fernandez-Borja *et al.*, 2005).

The RBD of mDia1 consists of several α-helices, but only two are prominently used in the interaction (Rose *et al.*, 2005) (Figure 1.11A). The structure of mDia1 with RhoC includes a series of hydrophobic and salt-bridge interactions between these helices and the two switches. Interestingly, mDia1 has a stabilising helix which helps the two main helices contact the switches but also makes minor contacts with RhoC. The RhoC insert also makes additional contacts with residues in a different domain of mDia1, which contains armadillo repeats.

1.3.2.2 The ROCK proteins

Rho-associated coiled coil kinase (ROCK) refers to two proteins, ROCK1 and ROCK2, which are RhoA effectors. ROCK is a serine/threonine kinase which phosphorylates the myosin light chain, leading to myosin-dependent actin cross-linking. ROCK also phosphorylates LIM-kinase, which in turn phosphorylates and inactivates the actin depolymerising factor, cofilin. Both of these events are important in the formation of actin structures such as stress fibres and it is unsurprising that ROCK is important in the assembly of the actomyosin ring during cytokinesis (reviewed in Narumiya *et al.*, 2009). ROCK is also thought to be critical for the invasion of cancer cells (Itoh *et al.*, 1999) while it has also been linked to apoptosis through its activation by caspase-3 (Sebbagh *et al.*, 2001).

ROCK1 homodimerises through its very long α-helical Rho-binding domain to form a parallel coiled coil which interacts with the switch regions of RhoA (Dvorsky *et al.*, 2004) (Figure 1.11A). This type of interaction is akin to those of other helical effectors interacting with the switch regions with parallel helices (reviewed in Mott & Owen, 2015). While all but one of the RhoA residues involved in the interaction are conserved in Rac1 and Cdc42, ROCK1 is an

exclusive effector of RhoA. This is thought to be due to a Glu residue which forms a salt bridge with a Lys of ROCK1. This Glu is an Asp in Rac1 and Cdc42 and may therefore be too short for the salt bridge to successfully form. Such subtle differences may be all that is needed for effectors to be specific to a particular GTPase, making it hard to predict specificity.

1.3.2.3 Citron kinase

Citron kinase (CIK) is a serine/threonine kinase which was first identified as a Rho effector involved in cytokinesis (reviewed in Thumkeo *et al.*, 2013; D'Avino, 2017). In addition to the kinase domain (which is absent from its neurone-specific isoform), CIK also contains two coiled-coil domains amongst others. The second coiled coil domain contains the RhoA-binding site, but the structure of this domain in complex with RhoA has not been determined.

RhoA in complex with CIK is part of the midbody complex of proteins which forms during cytokinesis. In particular, CIK is thought to serve as the link between the actomyosin contractile ring and the central spindle. In addition to its structural role, CIK is involved in a cross-regulation with Aurora B kinase. Phosphorylation of Aurora B kinase by CIK inactivates the former. Binding of RhoA is thought to prevent the inactivation of the Aurora B kinase which is required for the eventual disassembly of the contractile ring. In this sense, CIK appears to act as more of a regulator than an effector. Failure of cytokinesis as a result of CIK mutations is thought to cause microcephaly in both rodents and humans.

1.3.2.4 The Rhotekins and the Rhophilins

Rhophilin and Rhotekin are scaffold proteins which were identified on the basis of their Rho binding domain (RBD) which is an HR1 domain that is homologous to the HR1 domain of PRK1 (Watanabe *et al.*, 1996; Reid *et al.*, 1996). The physiological roles of the Rhophilins and the Rhotekins are poorly understood, although Rhophilin has primarily been linked to RhoA-dependent suppression of actin fibres while Rhotekin has been implicated in synapse regulation in neurones and outgrowth of axons (reviewed in Thumkeo *et al.*, 2013).

The Rhotekins (Rhotekin1 and Rhotekin2) have an N-terminal HR1 domain which has ~41% sequence identity to the HR1a domain of PRK1 and also contain a PH domain that is speculated to bind to lipids in the plasma membrane (reviewed in Ito *et al.*, 2018). The Rhotekin HR1 binds to RhoA with a higher affinity than PRK1 HR1a does (Blumenstein & Ahmadian, 2004), yet the reasons for this, whether structural or sequence-driven, have not been explored. In contrast, the Rhophilins (Rhophilin1 and Rhophilin2) contain BRO1 and PDZ protein-protein interaction domains in addition to their N-terminal HR1 domain. The affinity of the Rhophilin HR1 for RhoA has not been investigated to date.

1.3.2.5 The PRKs

The Protein Kinase C-related kinase (PRK) family consists of three members, PRK1, PRK2 and PRK3. PRK1 was the first identified effector of RhoA (Watanabe et *al.*, 1996). The HR1 domains function as the RBD of the PRKs (Flynn *et al.*, 1998; Owen *et al.*, 2003, Modha *et al.*, 2008, Hutchinson *et al.*, 2013).

1.4 The Protein Kinase C-related kinase family

1.4.1 Discovery

The PRK family of serine/threonine kinases consists of three members, PRK1 (also PKN, PKN1 and PKNα), PRK2 (also PKN2, PKNγ) and PRK3 (also PKN3, PKNβ). PRK1 was first identified from a human hippocampus cDNA library (Mukai and Ono, 1994) and was shown to have a catalytic domain related to that of the protein kinase C superfamily. It is found on chromosome 19 (Bartsch *et al.*, 1998). PRK2 (chromosome 1) and PRK3 (chromosome 9) were later identified as members of the same family due to their sequence similarity to PRK1 (Palmer *et al.*, 1995; Oishi *et al.*, 1999). PRK2 and PRK3 show 87% and 70% conservation to PRK1, respectively.

1.4.2 Domain arrangement

The domain arrangement of the PRKs is illustrated in the schematic in Figure 1.13.

1.4.2.1 The N-terminal HR1 domains

PRKs have three N-terminal HR1 domains, HR1a, HR1b and HR1c. They are predicted antiparallel coiled coils of approximately 10 kDa and together they form the regulatory region of the PRKs. The HR1 domains are also effector domains of Rho GTPases, with RhoA being the first GTPase identified as an interacting partner of PRK1, specifically interacting with HR1a (Watanabe *et al.*, 1996; Flynn *et al.*, 1998). Binding to Rho GTPases is how HR1 domains regulate PRK activity.

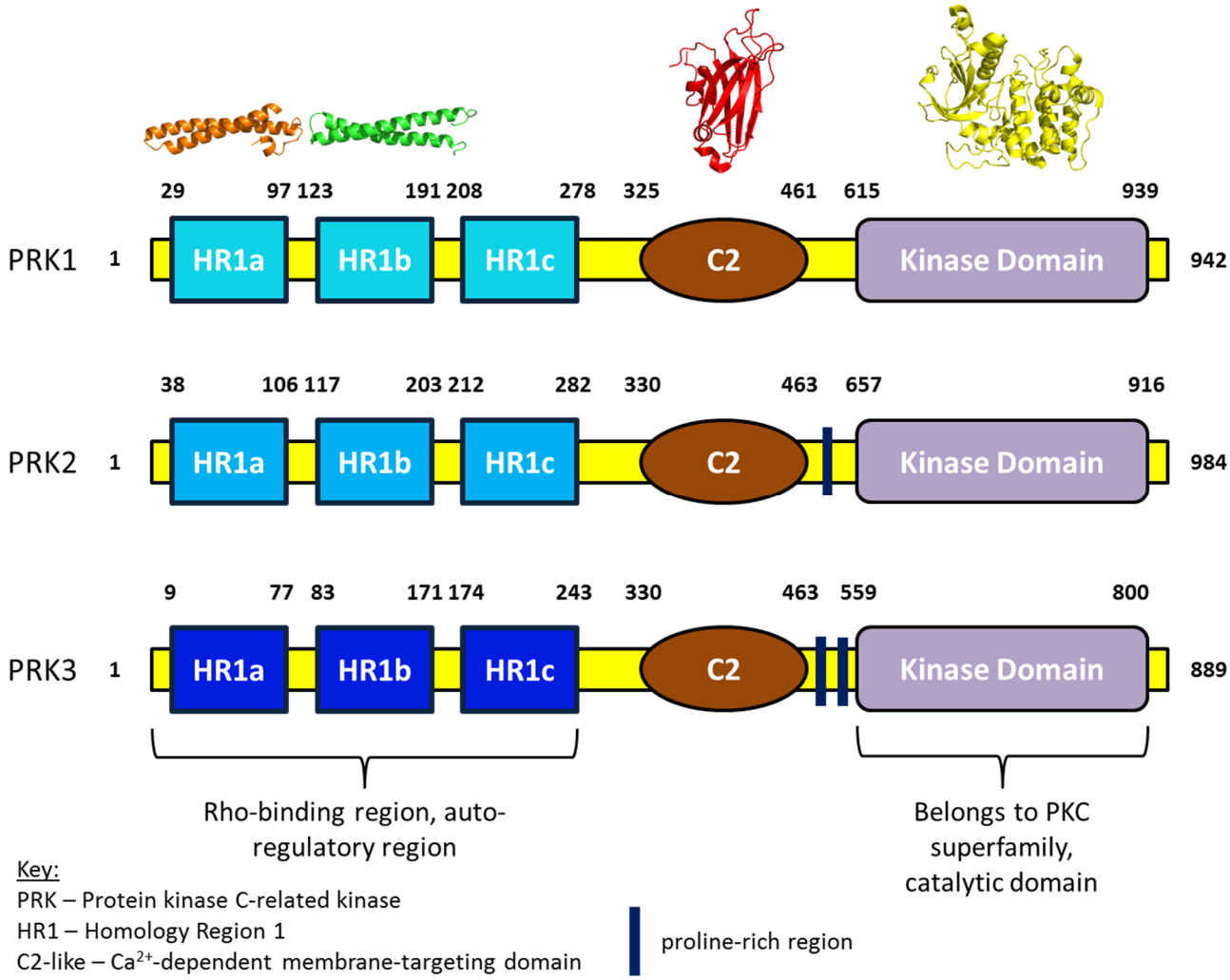


Figure 1.13 Domain structure of PRK proteins. The family has three members, PRK1, PRK2 and PRK3. Each carries 3 tandem N-terminal HR1 domains, a C2-like domain which is thought to target them to the membrane and a catalytic kinase domain which belongs to the PKC superfamily. In addition, PRK2 and PRK3 have one and two proline-rich regions, respectively. Solved structures of these domains, or an equivalent structure from another family isoform in the case of the C2-like domain are seen above the schematic; PRK1 HR1a (PDB:1CXZ), PRK1 HR1b (PDB:2RMK), C2-like domain from PKC (PDB:2UZP) and PRK1 catalytic domain (PDB:4OTD). No structure exists of an HR1c domain.

1.4.2.2 The C2-like domain

PRKs have a C2-like domain which shows weak homology to the C2 domains of other PKC proteins. C2 domains comprise 8 anti-parallel β-sheets which form a β-sandwich that binds to membrane phospholipids in a Ca^{2+}-dependent manner. They are therefore critical for the recruitment of proteins to the plasma membrane (reviewed in Corbalán-García & Gómez-Fernández, 2010). The C2-like domains of PRK are insensitive to Ca^{2+} but bind lipids to activate the PRKs (see sections 1.4.3 and 1.4.4).

1.4.2.3 Proline-rich regions

Quilliam *et al.* (1996) identified a proline-rich region in PRK2 and showed that it can interact with the middle SH3 domain of the NCK adaptor protein and also more weakly with the SH3 domain of phospholipase Cγ. The PRK2 amino acid motif DL**E**PE**PP**PA**P**P**R**AS**S**L fits the consensus X_2**E**X_2**PP**X_2**P**X**R**X_2**S**X which associates with the NCK SH3 domain. This proline-rich region is absent in PRK1. In contrast, PRK3 has two proline-rich regions, both of which fit the same consensus as the PRK2 motif and are also involved in SH3 domain-mediated protein-protein interactions (Shibata *et al.*, 2001; Oishi *et al.*, 1999; see section 1.4.5.9).

1.4.2.4 Kinase domain

The PRKs are members of the AGC kinase family, which includes the master protein kinases (PK) PKA, PKB, PKC and PKG amongst others. In particular, the PRKs have a C-terminal catalytic domain which belongs to the protein kinase C superfamily. In addition to the sequence homology of their catalytic domain, AGC kinases share three conserved features; firstly, a C-terminal extension of the catalytic domain which contains a hydrophobic motif (HM) that usually includes a phosphorylation site, secondly, a PDK1-interacting fragment (PIF) that binds to the activating loop kinase, PDK1, and thirdly, a turn motif (TM) (C-terminal to the catalytic domain) (reviewed in Pearce *et al.*, 2010; Arencibia *et al.*, 2013). Phosphorylated HM motifs recruit activating kinases like PDK1 and can interact with the PIF motif of either PDK1 or of their own molecule, forming an intramolecular interaction which allosterically activates the kinase (reviewed in Leroux *et al.*, 2018). PRKs have a phosphomimic residue at the HM phosphorylation site, so they are not regulated by phosphorylation of their HM. Phosphorylation of the TM has been reported to increase the affinity of the PRK2 C-terminal extension for the kinase domain to enhance activation (Dettori *et al.*, 2009).

The crystal structure of the PRK1 kinase domain has been solved to 2.0 Å by molecular replacement to PKCθ (Chamberlain *et al.*, 2014). The structure resembles AGC kinases with an N-lobe and C-lobe. Similar to other AGC kinases, e.g. PKA, the C-terminal extension of the PRK1 kinase domain encircles the catalytic domain and forms an active site tether, with

Phe910 docking into the active site of the kinase domain. This may provide the structural explanation of some of the activatory mechanisms that have been reported.

1.4.3 Activation

PRKs can be activated in several ways (Figure 1.14). PRK1 is the best-characterised of the PRK isoforms and it is assumed that PRK2 and PRK3 share at least some of its mechanisms of activation. Activated RhoA binds and activates PRK1 (Amano *et al.*, 1996; Lu & Settleman, 1999). Both PRK1 and PRK2 require an intact C-terminus for maximum activation by RhoA (Lim *et al.*, 2006; Lim *et al.*, 2008). Activation by RhoA is further supported by the *Drosophila* PRK which is activated by RhoA and Rac1 (Lu & Settleman, 1999). RhoC has also been linked to increased PRK3 activity (Unsal-Kacmaz *et al.*, 2012).

PRK1 can also be activated by phosphorylation of its activation loop by phosphoinositide-dependent protein kinase-1 (PDK1) (Dong *et al.*, 2000). This interaction was identified by a yeast two-hybrid screen using mouse PDK1 as bait against a cDNA library generated from HeLa cell mRNAs as prey. A co-immunoprecipitation from CHO cells confirmed this interaction. Deletion mutants further narrowed the interaction to be between the PDK1 and PRK1 catalytic domains, while phosphopeptide analysis identified several PRK1 Ser and Thr residues which were phosphorylated, primarily Thr774 (in activation loop). This phosphorylation was later confirmed in CHO/IR cells and shown to activate PRK1 activity, thereby confirming PDK1 as an *in vivo* activating kinase of PRK1. Additionally, PRK1 has to be in complex with RhoA for this activation (Flynn et *al.*, 2000). PRK2 has also been shown to be activated by PDK1 (Balendran *et al.*, 2000), and sequence analysis of PRK3 suggests that this third isoform of the family is regulated in an analogous manner. Specifically, mutation of the phosphoacceptor residue (in activation loop) in PRK2 has been shown to abrogate its activation, as did mutation of the TM phosphoacceptor residue, demonstrating that dual phosphorylation is required for activation (Lim *et al.*, 2008). The TM of PRK3 must also be phosphorylated for activation (Unsal-Kacmaz *et al.*, 2012).

PRK1 and PRK2 can also be activated by proteolysis, such as by caspase-3 during apoptosis (Mukai *et al.*, 1994; Cryns *et al.*, 1997; Takahashi *et al.*, 1998).

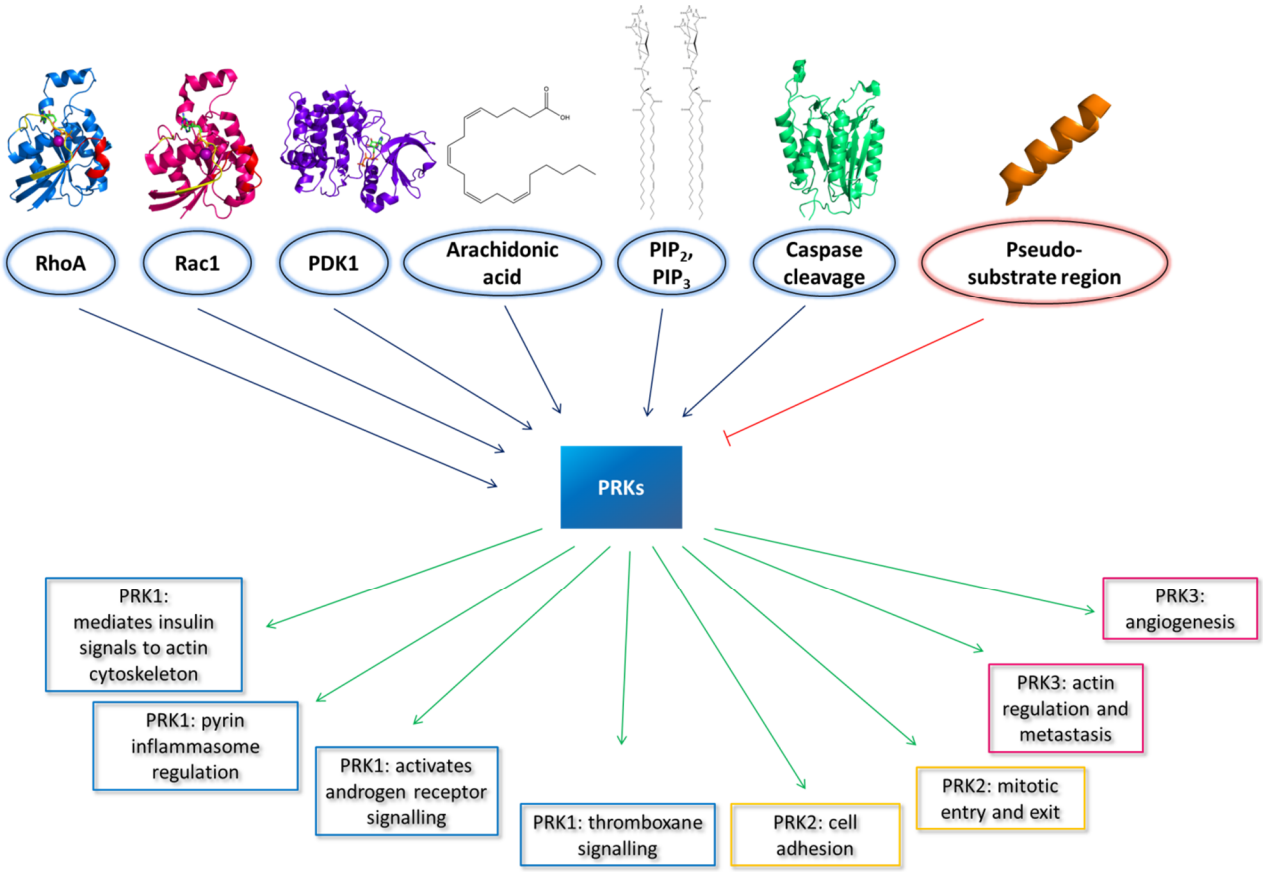


Figure 1.14 Activation of PRKs and downstream cellular effects. Blue arrows indicate PRK activation. The red arrow indicates PRK autoinhibition. Green arrows indicate key PRK cellular activities. PDK1 (phosphoinositide-dependent protein kinase-1), PIP_2 (phosphatidylinositol 4,5-bisphosphate), PIP_3 (phosphatidylinositol 3,4,5-triphosphate).

PRK1 has also been shown to be activated by phosphatidylinositol 4,5-bisphosphate (PIP_2) and phosphatidylinositol 3,4,5-triphosphate (PIP_3) *in vitro* (Palmer *et al.*, 1995). Finally, PRK1 can be activated by arachidonic acid (Mukai *et al.*, 1994), and the last 4 C-terminal residues appear to be required for this activation (Lim *et al.*, 2005). Arachidonic acid is thought to remove the autoinhibition exhibited by a region of PRK1 and PRK2 which overlaps with the C2-like domain (Yoshinaga *et al.*, 1999). Arachidonic acid also activates PRK3; however activation of both PRK2 and PRK3 by arachidonic acid is less than that of PRK1 (Yu *et al.*, 1997; Oishi *et al.*, 1999). The activation is associated with an increase in autophosphorylation. In addition to arachidonic acid, PRKs are activated by additional lipids which include linoleic acid, oleic acid, linolenic acid as well as cardiolipin which leads to greater increase in PRK2 than PRK1 phosphorylation (Morrice *et al.*, 1994; Kitagawa *et al.*, 1995; Peng *et al.*, 1996; Yu *et al.*, 1997; Mukai & Ono, 2006). Furthermore, the C-terminus of PRK2 may prevent activation by cardiolipin (Lim *et al.*, 2008). A kinetic study has recently confirmed that the PRKs show distinct lipid activation and that the mechanisms of activation may differ as indicated by arachidonic acid increasing the k_{cat} of PRK1 and PRK2 but decreasing the K_m of PRK3 (Falk *et al.*, 2014). Differential activation by lipids may signify distinct temporal activation of the PRKs or that they get activated in different compartments, the membranes of which are enriched in different lipids.

Figure 1.14 summarises the key activatory molecules of PRKs and some of the downstream effects.

1.4.4 Autoregulation

Like other kinases, PRKs are involved in a number of autoregulatory mechanisms. Kitagawa et *al.* (1996) first reported evidence of this when they showed that a peptide with sequence corresponding to the amino acids 39-53 of PRK1 (part of the HR1a domain), and with an isoleucine mutated to a serine, could be phosphorylated by wild type PRK1. The wild type peptide did not get phosphorylated. This suggests that HR1a contains a pseudosubstrate region which can bind to the kinase domain and keep the PRK in an autoinhibited state. This inhibition is thought to be relieved when RhoA binds to HR1a, freeing up the kinase domain. Observations consistent with this regulation of PRK1 have been made in cells (Manser *et al.*, 2008).

Yoshinaga *et al.* (1999) also reported that a region of PRK1 and PRK2 which overlaps with the C2 domain is involved in autoinhibition. This autoinhibition was found to be relieved by arachidonic acid, highlighting the need of the PRKs to be recruited to a membrane in order to be successfully activated.

Consistent with the autoinhibition of PRK1, the N-terminus of PRK2 has recently been shown to both inhibit its catalytic domain as well as its interaction with PDK1 (Bauer *et al.*, 2012). Specifically, PRK2 activity appears to be dependent on the concentration-dependent oligomerisation of the molecule. Cross-linking experiments with His-tagged purified full length PRK2 stabilised dimers and tetramers, while similar experiments with FLAG-PRK2 purified from HEK 293T cells stabilised only dimer, suggesting that dimer and not tetramer is the physiological oligomer which forms *in vivo*. Using deletion mutants, the authors were able to define this region of intermolecular PRK2-PRK2 interactions as amino acids 464-500, which is located between the C2 domain and the catalytic domain.

Kinases often autophosphorylate, resulting in the activation or inhibition of their catalytic activity. The PRKs have also been shown to autophosphorylate in response to lipids and several phosphorylation sites have been identified throughout the molecule (Peng *et al.*, 1996; Oishi *et al.*, 1999; Yoshinaga *et al.*, 1999). Autophosphorylation was shown to activate the catalytic activity of the PRKs.

1.4.5 Role in cellular signalling and physiological roles

Table 1.2 lists some of the key PRK substrates which have been identified. The cellular processes in which PRKs are involved will be discussed below.

PRK	Substrate	Cell process/Function
PRK1	α-actinin	Cytoskeletal regulation
PRK1	vimentin	Cytoskeletal regulation
PRK1	Tau	Alzheimer's disease
PRK1	GSK-3β	Glucose metabolism
PRK1	α Crystallin B	Molecular chaperone
PRK1, PRK2	HDAC5 (Thr292), HDAC7, HDAC9	Chromatin remodelling
PRK1	H3 Thr11	Androgen receptor signalling
PRK1	ARHGAP15	RhoGAP
PRK1, PRK2	Cdc25	Cell cycle
PRK1, PRK2	Pyrin	Immune response
PRK1	EGFR	Cell signalling
PRK1, PRK3	Phospholipase D1	Cell signalling
PRK2	Cortactin	Cytoskeletal rearrangement
PRK2	DUSP6	MAPK cascade

Table 1.2 Key substrates of the PRKs. The PRK isoform, substrate and corresponding cellular process/function are listed.

1.4.5.1 Cytoskeletal regulation and cell adhesion

The PRKs are widely involved in cytoskeletal regulation. Mukai *et al.* (1997) were the first to suggest this when they performed a yeast two-hybrid screen which identified α-actinin, an actin-binding protein , as an interacting partner of PRK1. In addition, overexpression of PRK1 in fibroblasts resulted in actin reorganisation, while PRK2 only had a small effect (Lim *et al.*, 2004). PRK1 has been shown to disrupt neurofilament assembly and axonal transport (Manser *et al.*, 2008). Furthermore, PRK1 was found to bind and phosphorylate the intermediate filament protein vimentin, leading to the disruption of vimentin filaments *in vitro* (Matsuzawa *et al.*, 1997). Knock-down of PRK3 has also been shown to lead to vimentin disruption in an orthotopic mouse prostate tumour model (Unsal-Kacmaz *et al.*, 2012). The actin cross-linking protein α-actinin is another substrate of PRK1 and its affinity for PRK1 is increased in the presence of PIP_2, a lipid which is involved in actin reorganisation (Mukai *et al.*, 1997).

The PRKs are also implicated in stress fibre formation. In particular, PRK1 was shown to be involved in insulin-dependent signalling to the actin cytoskeleton. Expression of PDK1 and PRK1 in 3T3-L1 and Rat1 cells led to insulin-induced actin cytoskeletal reorganisation and specifically to a reduction in the number of stress fibres as well as membrane ruffling. This process was prevented when the expressed PRK1 or PDK1 were catalytically inactive (Dong *et al.*, 2000). Particularly, expression of a kinase-dead mutant of PRK2 led to almost complete loss of actin stress fibres, while expression of wild type PRK2 caused a slight reduction in the number of stress fibres while also leading into formation of larger fibres (Vincent & Settleman, 1997). Microinjection of dominant-negative PRK2 completely disrupted actin stress fibres (Lu & Settleman, 1999). Put together, the above studies suggest that both PRK1 and PRK2 may have a role in the disruption of actin stress fibres and the cellular reorganisation of F-actin.

PRK3 also appears to hold multiple roles in cytoskeletal organisation. RNAi-mediated knockdown of PRK3 expression in human umbilical endothelial cells (HUVECs) compromised these cells' ability to migrate (Aleku *et al.*, 2008) and also reduced their ability to form stress fibres downstream of TNF-α signalling (Möpert *et al.*, 2012). The latter study also linked PRK3 to cell adhesion, as PRK3-depleted HUVECs showed irregular VE-catenin and β-catenin distribution, proteins which are part of adherens junctions. Furthermore, PRK3-depletion also led to a decrease in expression levels of the ICAM-1 glycoprotein and a decrease in the phosphorylation of Pyk2 tyrosine kinase, both of which are involved in cell adhesion, suggesting that PRK3 either has a direct role in cell adhesion or its loss leads to actin cytoskeletal changes which in turn affect cell adhesion. A more recent study has implicated PRK3 in the glycosylation of ICAM-1 and integrins β1 and α5 through the regulation of N-acetylglucosaminyltransferase III and glycosyltransferase, suggesting that PRK3 may indeed have a direct role in cell adhesion (Mukai *et al.*, 2016).

Bourguignon *et al.* (2007) examined astrocyte activation and migration and determined that these processes are dependent on Rac1-dependent PRK2 activation and specifically the phosphorylation of cortactin. Cortactin is a cytoskeletal protein which cross-links F-actin and its phosphorylation by PRK2 inhibits this process, leading to important cytoskeletal rearrangements which are key to astrocyte migration. Expression of the PRK2 HR1 region acted as a dominant negative, reducing cortactin phosphorylation and preventing

cytoskeletal rearrangements, thus confirming that Rac1 activates PRK2 activity by binding to the HR1 domains.

Thrombin is a serine protease which causes endothelial cell permeability. Gavard & Gutkind (2008) showed that thrombin-dependent activation of endothelial cell protease-activated receptor 1 (PAR1) activates the heterotrimeric G proteins $G_{\alpha 12/13}$ and $G_{\alpha 11/q}$. These G proteins then activate RhoA which in turn activates ROCK to stimulate actomyosin contractility. PRK1 was also shown to be activated leading to the disruption of cell-cell adhesion, resulting in the ROCK/PRK1 synergistic stimulation of vascular leakage.

Calautti *et al.* (2002) investigated the activation of Fyn tyrosine kinase in differentiated keratinocytes. They found that increased Fyn tyrosine kinase activity was linked to increased RhoA expression. They also discovered that PRK2 induces downstream, Fyn/Src-dependent tyrosine phosphorylation of β-catenin, γ-catenin and $p120^{ctn}$, which promote keratinocyte cell-cell adhesion. PRK2 therefore directly or indirectly activates Fyn and Src tyrosine kinases. The role of PRK2 in cell adhesion has been further demonstrated in C2C12 myoblasts (Lee *et al.*, 2016). In this study, PRK2 was found to interact directly via its C-terminal region with Cdo, a cell-surface protein, PKB/Akt and the adaptor protein APPL1, leading to PKB/Akt activation which promotes the differentiation of myoblasts. This is in contrast to PRK1 which has been shown to be a negative regulator of PKB/Akt downstream of B-cell antigen receptors (Yasui *et al.*, 2012).

In neurons affected with Alzheimer's disease, PRK1 has been linked to neurofibrillary tangles and their key component, the protein tau (Kawamata *et al.*, 1998). In particular, PRK1 phosphorylates the tau protein which is thought to induce key conformational changes (Isagawa *et al.*, 2000). GSK-3β also phosphorylates the tau protein. However it was shown that phosphorylation of GSK-3β by PRK1 reduced GSK-3β-dependent phosphorylation of tau. PRK1 is also activated in neurons under hypoxia and its loss has been linked to decreased cell viability and disrupted actin cytoskeletal organisation (Thauerer *et al.*, 2014). All of the above implicate deregulation of PRK1 in Alzheimer's disease.

1.4.5.2 Androgen receptor signalling, chromatin remodelling and transcriptional regulation

Metzger *et al.* (2003) implicated PRK1 and PRK2 in androgen receptor (AR) signalling. They showed that PRK1 binds to the transactivation domain of the AR in vitro and also confirmed the interaction to occur *in vivo* by co-immunoprecipitation experiments. PRK1 and the AR bind together to the androgen response elements of AR-regulated genes. A subsequent study by Metzger *et al.* (2008) showed that PRK1 phosphorylates Thr 11 of histone H3 which is an epigenetic marker linked to the expression of genes regulated by the androgen receptor. H3 Thr11 phosphorylation recruits the demethylase JMJD2C which subsequently leads to demethylation of H3 Lys9 (methylated H3 Lys9 is associated with a repressive chromatin state). Mechanistically this explains the critical role of PRK1 in AR signalling and implicates it in prostate cancer.

Additionally, kinase assays have shown that both PRK1 and PRK2 bind and phosphorylate Thr292 of HDAC5 (K_M ~ 20 μM) which is within its nuclear localisation signal (NLS) (Harrison *et al.*, 2010). HDAC7 and HDAC9 are also phosphorylated by the PRKs in a similar manner. This phosphorylation is thought to prevent nuclear import by allowing the binding of 14-3-3 proteins to HDACs, therefore highlighting a direct effect the PRKs have on chromatin structure and as a result transcriptional regulation.

Rho family-dependent PRK2 activity was also shown to be sufficient for the transcriptional activation of the c-fos serum response element (Qulliam *et al.*, 1996). PRK1 was also shown to interact with a neurone-specific helix-loop-helix transcription factor, NDRF2, by yeast two-hybrid and co-IPs and to amplify NDRF2-mediated transcription (Shibata *et al.*, 1999). This further supports the notion that the PRKs may regulate transcription of currently undetermined genes.

1.4.5.3 Cell cycle

The PRKs are also implicated in the cell cycle (Schmidt *et al.*, 2007). PRK2, and to a lesser extent PRK1 (suggesting partial redundancy in function), were detected in the cytosol of HeLa cells during interphase but at the cleavage furrow during telophase and at the midbody during cytokinesis. Furthermore, siRNA-depletion of PRK2 led to failed abscission of mother and daughter cells, suggesting a role of the PRKs in the cytoskeletal

reorganisation which occurs at cytokinesis. In addition to the role of PRK2 in mitotic exit, it was also shown to be important in the G2 to M-phase transition. Cells with depleted PRK2 exhibited delayed mitotic entry and this phenotype was shown to correlate with reduced phospho-histone3 levels and delayed cyclin A/B1 degradation, which are known mitotic phase markers. Moreover, reduced phosphorylation of Cdc25 phosphatase was detected in PRK2-depleted cells, suggesting that Cdc25 is a substrate of PRK2 and is regulated by it. PRK2, as well as PRK1, are substrates of active Cyclin-Dependent Kinase (CDK) 1, suggesting a positive feedback loop which further activates the PRKs via Cdc25. In summary, PRK2 (and PRK1) are involved in both mitotic entry and exit. In contrast to the above, a previous study had shown that addition of active PRK1 into a *Xenopus laevis* embryo by microinjection caused cleavage arrest and delayed mitotic entry (Misaki *et al.*, 2001) and this was linked to the *in vitro* phosphorylation of Cdc25 by PRK1. This suggests that the PRKs may have opposite effects in the cells of different species, at least with respect to PRK1.

The link between the PRKs and CDKs/cyclins extends to the cyclinD1-CDK6 complex, which was found to phosphorylate PRK1 and modulate vascular smooth muscle cell division and migration which is important in blood vessel wall remodelling (Singh *et al.*, 2012). While this suggests that PRK1 is activated downstream of cyclinD1, a more recent study suggests that it is also located upstream of it (Wang *et al.*, 2017). In this study, RhoA was shown to transcriptionally regulate cyclinD1 expression via PRK1 in human embryonic stem cells leading to cell proliferation. Put together, the evidence from these two studies suggests a possible feedback loop between PRK1 and cyclinD1.

1.4.5.4 Immune response

The PRKs have recently been implicated in signalling involving the pyrin inflammasome (Park et *al.*, 2016). Pyrin is a protein encoded by the *MEFV* gene and it is involved in the formation of a protein complex known as the pyrin inflammasome. This inflammasome forms as part of the immune system's response to bacterial infection and one of its roles is to induce interleukin (IL) release from macrophages to combat infection. In this study, siRNA depletion in bone marrow-derived macrophages of PRK1 and PRK2 (but not of ROCK1 or ROCK2) led to inflammasome-dependent IL-1β release, suggesting that these two PRKs inhibit pyrin

inflammasome activation. In addition, the study showed that mouse pyrin binds directly to the PRK kinase domain, leading to phosphorylation of two serines. These phosphoserines recruit 14-3-3 scaffold proteins to inhibit inflammasome activation. Toxins such as the C3 toxin from *Clostridium botulinum* and TcdB from *Clostridium difficile* lead to the constitutive activation of RhoA and thus to PRK-mediated pyrin inflammasome inhibition. This study shows that inhibiting any component of the RhoA/PRK1 pathway is sufficient for pyrin inflammasome activation. The *MEFV* gene is mutated in patients with the autoinflammatory diseases Familial Mediterranean Fever and hyperimmunoglobulinemia D syndrome (HIDS). The authors knocked in a mutant form of human pyrin in mice (the mutation is in the B30.2 domain of pyrin which is exclusive to humans) and observed decreased binding to the PRKs, suggesting that the B30.2 has a role in the binding of PRK1 to pyrin. Finally, they observed spontaneous activation of the pyrin inflammasome pathway in peripheral blood mononuclear cells (PBMCs) from patients with HIDS, a result of a loss-of-function mutation in the *MVK* gene, which leads to the abolishment of RhoA geranyl-geranylation and as a result reduced PRK1 activation. Unsurprisingly, pyrin inflammasome activation in PBMCs was blocked with arachidonic acid. To summarise, PRK1 downstream of RhoA has a novel role in the downregulation of pyrin inflammasome formation, which implicates it in the innate immune response.

PRK1 has also recently been implicated in lymphocyte trafficking (Mashud *et al.*, 2017). Mouse lymphocytes with PRK1 T778A (in activation loop), an inactive PRK1 mutant, displayed reduced migration from secondary lymphoid organs to the blood and lymph, leading to a decreased number of lymphocytes in peripheral blood. This may be due to a severely decreased migratory response to sphingosine 1-phosphate which activates the S1PR GPCR, suggesting that PRK1 is implicated in a signalling pathway downstream of the receptor.

1.4.5.5 Thromboxane signalling

The Thromboxane receptors TPα and TPβ are GPCR receptors which are activated by thromboxane A_2 (TXA_2) produced by the arachidonic acid pathway. These GPCRs have been implicated in activation of RhoA via the G_q heterotrimeric G protein (Wikström *et al.*, 2008).

Turner *et al.* (2011) showed that PRK1 interacts with both TPα and TPβ and that this interaction is important for TPα- and TPβ-induced cell migration. All 3 PRKs were found to interact with the TP receptors in HEK293 cells but in PC-3 cells (an adenocarcinoma, prostate cancer cell line), only PRK1 and PRK2 were shown to interact (O'Sullivan *et al.*, 2015). Furthermore, there was differential association of the PRKs with the TP proteins, e.g. PRK1 and PRK3 constitutively formed a complex with both TPα and TPβ, while PRK2 constitutively associated with TPβ and only transiently associated with TPα (in an agonist-dependent manner). This interaction led to activation of PRK1 and PRK2 via phosphorylation of their turn motif in a PI3-K/PDK1-dependent mechanism. Activation of PRK1 and PRK2 in thromboxane signalling was found to be crucial for TXA_2-induced chromatin remodelling through phosphorylation of H3T11 but also enhanced recruitment of the androgen receptor which can also lead to H3T11 phosphorylation (O'Sullivan *et al.*, 2017). Finally, PRK1 and PRK2 were found to promote the migration and cellular proliferation of prostate cancer cells. To summarise, TP receptor activation leads to activation of G_q and subsequently RhoA, which is thought to recruit the PRKs which in turn can interact with the TP receptor to mediate downstream signalling. In this way, PRK1 and PRK2 are implicated in TP-mediated, androgen-induced cellular responses which are deregulated in cancer.

1.4.5.6 The role of PRK3 in angiogenesis

PRK3 is unique from the other PRKs in that it has a physiological role in angiogenesis (Mukai *et al.*, 2016). *Ex vivo* aortic ring and *in vivo* corneal pocket assays of PRK3 knockout mice showed decreased micro-vessel sprouting, while MEFs from these mice displayed reduced migration which could not be rescued even in the presence of growth factors. Surprisingly, PRK3 did not affect tumour angiogenesis, despite the observed reduction in lung metastasis of melanoma cells injected in the tail of the PRK3 knockout mice.

1.4.5.7 Apoptosis

The PRKs are activated during apoptosis by caspase cleavage. Cell extracts from COS-7 cells expressing FLAG-PRK1 were incubated with caspase-3 which resulted in the cleavage of PRK1 at more than one site (Takahashi *et al.*, 1998). PRK2 is also activated by cleavage

(Cryns *et al.*, 1997, Takahashi *et al.*, 1998) and the cleaved PRK2 has been shown to interact with PKB/Akt. The constitutively active PRK1 fragment has been suggested not to have a direct role in apoptotic events, but rather it may be involved in a different role, such as a mild change in morphology observed in microglia (Ueyama *et al.*, 2001). Furthermore, glutamate can induce caspase cleavage of PRK1 in neurons to generate the constitutively active kinase, which is linked to neurofilament organisation disruption (Manser *et al.*, 2008).

1.4.5.8 Glucose metabolism

PRK1 has an additional role in the regulation of glucose metabolism. The study by Dong *et al.* (2000) which identified the interaction of PRK1 with PDK1 (see section 1.4.3) also observed actin cytoskeletal reorganisation and membrane ruffling in 3T3-L1 fibroblasts and also in a Rat1 stable cell line expressing the insulin receptor (Rat1-IR) when the cells were treated with insulin or had PRK1 or PDK1 expressed. The inability of cells with non-phosphorylatable PRK1 or kinase-dead PDK1 to show equivalent cytoskeletal and membrane changes suggests that activation of PRK1 via PDK1 phosphorylation is vital for the signal transduction from the insulin receptor to the actin cytoskeleton.

In addition, the effect of PRK1 on insulin-stimulated glucose transport was shown to be dependent on both active RhoA and phosphoinositide-3-kinase (PI3-K) (both upstream of PRK1), as suggested by a study on rat adipocytes stimulated with GTPγS or insulin (Standaert *et al.*, 1998). To some extent this should be unsurprising, as both RhoA and phosphoinositides produced by PI3-K-catalysed reactions are known to activate PRK1 (see section 1.4.3). Transient expression of constitutively active RhoA and wild type PRK1 also increased the levels of GLUT4 glucose transporters translocated to the membrane. This translocation was abolished in the case of dominant negative RhoA or PRK1.

PRK1 has further been implicated in glucose metabolism and specifically glycogen synthesis. It has been shown that PRK1 can phosphorylate and inhibit glycogen synthase kinase, GSK-3β, an enzyme which inactivates glycogen synthase (Isagawa *et al.*, 2000).

1.4.5.9 Interactions with other signalling proteins

The PRK3 proline-rich region was used as bait against a HEK293 cDNA library in a yeast two-hybrid assay and Graf2, an SH3-containing protein also with a RhoGAP domain for RhoA and Cdc42, was identified as an interacting partner (Shibata *et al.*, 2001). *In vitro* experiments with recombinant proteins also showed that PRK3 interacts with both Graf and Graf2. The significance of these findings is that PRK3 may be unique amongst the PRK isoforms by recruiting a RhoGAP protein, bringing it into proximity with Rho GTPases to terminate signalling.

The PRK2 proline-rich region was instead found to interact with the SH3 domain of the NCK adaptor protein and also weakly with the SH3 of phospholipase Cγ (Quilliam *et al.*, 1996), as well as one of the SH3 domains of NCKβ (also known as Grb4) (Braverman & Quilliam, 1999). NCK interacts with receptor tyrosine kinases via its SH2 domain while NCKβ causes cellular transformation through the v-Abl tyrosine kinase (a proto-oncogene). PRK2 therefore may act as a direct mediator in signalling downstream of activated receptor tyrosine kinases. The interaction of NCK with other Rho family effectors such as WASP also highlights possible cross-talk between Rho effector pathways.

PRK2 also binds to MEKK2, a mitogen-activated protein kinase (MAPK) kinase kinase (Sun *et al.*, 2000). The binding region is located just N-terminal of the catalytic domain (residues 637-660). In contrast, PRK1 did not bind to MEKK2 (residues in the binding region are poorly conserved). MEKK2 was found to activate PRK2 directly, suggesting that signalling downstream of a MAPK pathway can activate a specific PRK isoform, PRK2.

Oishi *et al.* (2001) investigated the co-localisation of the PRK family with phospholipase D1 and showed that PRK1 and PRK3 interact with this protein by co-immunoprecipitations from COS-7 cells. They then confirmed the interaction *in vitro* and showed that PRK1 binds to the internal region (residues 298-598) of phospholipase D1, while PRK3 binds to both the N-terminal (residues 1-228) and the internal regions. This differential binding of the PRKs may in turn explain the higher affinity of phospholipase D1 for PRK3 (due to cooperative binding to two of its regions) but it may also explain the differential activation of the phospholipase D1 catalytic activity by the PRKs (greater activation for PRK1 than PRK3). Phospholipase D1 is a regulator of several cell processes, including differentiation, GLUT4 vesicle trafficking,

cytoskeletal reorganisation and cell growth (Banno, 2002), therefore the PRKs can be considered as upstream regulators of these processes.

H_2O_2-mediated activation of PRK1 by phosphorylation of its kinase loop threonine has been linked to the survival of cardiac myocytes, direct phosphorylation of α crystallin B (a molecular chaperone) and increased cardiac proteasome activity, events suggesting protection of heart during ischaemia (Takagi *et al.*, 2010). Hypotonic swelling of cardiac myocytes also activates PRK1 and may involve signalling between Src, RhoA, PRK1 and ERK which leads to survival (Kajimoto *et al.*, 2011).

The mammalian target of rapamycin (mTOR), a protein complex with diverse signalling roles, including regulation of the cytoskeleton (Jacinto *et al.*, 2004), has recently been shown to be critical for the phosphorylation of the turn-motif of PRK1, which leads to its activation (Yang *et al.*, 2017). This further implicates the PRKs as effectors downstream of PI3K in cellular signalling.

Several other substrates of PRK1 have been identified, notably ARHGAP15 (a Rho GAP), the Ret kinase (a proto-oncogene) and exonuclease 1 (implicated in DNA repair), suggesting the existence of many cellular roles of the PRKs which remain uncharacterised (Collazos *et al.*, 2011). More recently, kinase-interacting substrate screening has identified several other possible candidate substrates for PRK1, including ROCK2, which could suggest cross-talk between RhoA effector-mediated pathways, and Cdc42EP4 (a Cdc42 and TC10 effector protein with a role in cytoskeletal organisation), suggesting cross-talk between effectors of distinct Rho subfamilies (Amano *et al.*, 2015). The best PRK1 and PRK3 substrate recognition consensus motif appears to be Lys-Arg-Arg-Lys-Pro-Ser-Phe-Arg-Asn-Pro and is similar to the consensus phosphorylation motif of PKC (Shiga *et al.*, 2010; Collazos *et al.*, 2011).

1.4.5.10 Development

Out of the three PRKs, PRK2 has specifically been linked to development. While PRK1 and PRK3 were shown to be dispensable during mouse embryogenesis, PRK2 was instead found to be essential, as indicated by PRK2 knock-out embryos which showed severe cardiovascular abnormalities, a collapsed mesenchyme and a neural crest migration defect,

suggesting a role for PRK2 in mesenchymal growth and neural crest migration *in vivo* (Quétier *et al.*, 2016). This is also supported by PRK2 knock-out embryos, which show insufficient axial turning and neural tube closure (Danno *et al.*, 2017), effects which were unobserved in the case of PRK1 and PRK3 knock-out embryos. Mouse embryonic fibroblasts from PRK2 knockout embryos displayed abnormalities in terms of cell proliferation, consistent with PRK2's role in the cell cycle. Out of the three PRKs, PRK2 is therefore the isoform which has a distinct role in development. This view is further strengthened by evidence of PRK2 regulating apical junction formation in bronchial endothelial cells (Wallace *et al.*, 2011). In particular, PRK2 is activated by RhoA and recruited to primordial junctions in a RhoA-dependent manner. A dominant negative (kinase-dead) isoform of PRK2 was unable to lead to the maturation of primordial to apical junctions, as was a ΔC2 mutant, suggesting that localisation via the C2 domain and also the catalytic activity are both indispensable for epithelial morphogenesis.

Recently, PRK1 has been identified as a key player in cerebellar development (zur Nedden *et al.*, 2018). PRK1 knockout mice had problems with axonal outgrowth and also with forming synapses in Purkinje cells, suggesting that PRK1 is required for these developmental events. PRK1 may exert these effects by regulating Akt activity, at least under this particular biological context. A more recent study in Drosophila indicated that the *Drosophila melanogaster* isoform of PRK1, Pkn, affects the spatiotemporal regulation of myosin, leading to its relocalisation, an event which is thought to promote the asymmetric division of neural stem cells. A mutation in the *Drosophila Pkn* gene, termed the *delorean* mutation, has also been linked to wing morphogenesis in the context of the Jun-terminal kinase pathway (Sass & Ostrow, 2014), while Pkn activity has also been linked to negative regulation of actin-myosin contraction in other developmental processes (Ferreira *et al.*, 2014).

1.4.6 Cellular localisation and tissue-specific expression

1.4.6.1 Cellular localisation

PRK1 and PRK2 have been detected in the cytosol (Gross *et al.*, 2001; Schmidt *et al.*, 2007; Unsal-Kacmaz *et al.*, 2012) while rat PRK1 has also been detected at the plasma membrane

(Zhu *et al.*, 2004). The PRKs are all thought to be activated by RhoA and PDK1 which are bound to the plasma membrane, so they most likely all localise to that compartment via their C2-like domain.

PRK1 was shown to co-localise with RhoB in the endosomes of Swiss 3T3 cells both by immunofluorescence and subcellular fractionation (Mellor *et al.*, 1998), suggesting that RhoB targets PRK1 to this compartment. This co-localisation required active RhoB and is dependent on the HR1 region of PRK1 as indicated by RhoB Q63L (active mutant) and HR1 deletion mutants, respectively. The ΔHR1 mutant failed to localise to the endosomes. This is consistent with RhoB interacting with HR1 domains in a GTP-dependent manner. In another study, co-expression of PRK1 with RhoB showed that RhoB regulates epidermal growth factor receptor (EGFR) trafficking through endosomes via PRK1 (Gampel *et al.*, 1999). This was disrupted when a kinase-dead mutant of PRK1 (K644M) was used, indicating that RhoB-mediated localisation in endosomes has a functional role. This may be to trap EGFR in late endosomes, preventing their trafficking to lysosomes and their eventual degradation which would therefore prolong EGFR signalling (Sandilands & Frame, 2008).

The PRKs have been detected in several other cellular compartments. PRK1 (Metzger *et al.*, 2003), PRK2 (Gross *et al.*, 2001) and PRK3 (Oishi *et al.*, 1999; Unsal-Kacmaz *et al.*, 2012) have all been detected in the nucleus, suggesting that they contain an NLS or are imported into the nucleus in complex with another protein. PRK1 has also been detected at the cell midbody and cleavage furrow of dividing cells, as has PRK2 (Schmidt *et al.*, 2007). Additionally, PRK2 has been detected on membranes in lamellipodia (Gross *et al.*, 2001) and in apical junctions (Wallace *et al.*, 2011). In contrast, PRK3 has also been detected in the perinuclear Golgi apparatus (Oishi *et al.*, 1999; Unsal-Kacmaz *et al.*, 2012).

The diversity in the cellular locations of the PRKs is not unexpected based on their diverse cellular functions. Moreover, the differential localisation of the PRKs may partly explain their lack of functional redundancy. Finally, it is important to keep in mind that the studies mentioned above investigated PRK localisation in specific cell types or under specific conditions, e.g. when RhoB is active. This suggests that the PRKs may localise to other compartments which are currently unknown, but more importantly that there is dynamic trafficking of the PRKs between the various compartments. This trafficking may be mediated

by the Rho GTPases themselves, e.g. RhoB/PRK1 in endosomes (Mellor *et al.*, 1998); RhoA/PRK3 in perinuclear regions and RhoC/PRK3 in the nucleus (Unsal-Kacmaz *et al.*, 2012). Another example would be the Rac1- and PDK1-dependent localisation of PRK1 on the surface of vesicles under hyperosmotic stress (Torbett *et al.*, 2003).

1.4.6.2 Tissue-specific expression

PRK1 and PRK2 show ubiquitous tissue expression (Kitagawa *et al.*, 1995; Qulliam *et al.*, 1996; Vincent & Settleman, 1997), whereas PRK3 was initially undetected in normal tissues with northern blotting, but was instead detected in metastatic cancers (Oishi *et al.*, 1999). More recently, a study using qPCR which is more sensitive was able to detect PRK3 mRNA in adult mouse tissues (Mukai *et al.*, 2016), suggesting that PRK3 is present in normal cells and that it does indeed have a physiological role. Interestingly, PRK3 mRNA levels were more than threefold higher in testes than in other tissues such as lung, kidney, spleen and heart. Brain and skeletal muscle tissues contained the lowest levels of PRK3 mRNA. This is in contrast to PRK1, which is highly expressed in brain tissue, where it is estimated that it consists of 0.01% of total brain protein (Kitagawa *et al.*, 1995). Therefore, PRK isoform-specific activity is also regulated at the level of differential tissue expression.

1.4.7 Role in disease

The PRKs are involved in diseases such as cancer or those caused by pathogens. The lack of redundancy between the PRKs is highlighted once again by their differential involvement in these pathologies, suggesting that to some extent they are involved in distinct molecular events and pathways. The PRKs possess astonishingly diverse cellular roles; this may explain their prominence in an equally diverse set of pathologies.

1.4.7.1 Cancer

The first indication that a PRK may be involved in cancer came from Gao *et al.* (2000). A yeast two-hybrid screen identified PRK1 as an interacting partner of the E6 oncoprotein of human papilloma virus (HPV). The E6 oncoprotein is known to inactivate the p53 tumour suppressor by targeting it for ubiquitination and proteosomal degradation. PRK1 was confirmed to interact with the E6 oncoprotein *in vivo* and to also phosphorylate it, which may lead to E6-induced cell immortalisation.

PRK1 is overexpressed in human prostate cancer and specifically co-expressed with the androgen receptor *in vivo* (Metzger *et al.*, 2003). As discussed previously, PRK1 promotes the transcription of androgen receptor-regulated genes by phosphorylating histone H3 Thr11. This is considered to be an epigenetic marker of prostate cancer (Metzger *et al.*, 2008). The importance of PRK1 in prostate cancer is further emphasized by the inhibition of PRK1 leading to abrogated cellular proliferation of AR-induced tumour cells. Furthermore, PRK1 and PRK2, activated downstream of thromboxane receptors, have been shown to both mimic and enhance androgen-dependent chromatin remodelling, also in the context of prostate cancer (O'Sullivan *et al.*, 2017). PRK1 therefore appears to be a key driver of prostate cancer. PRK1 mRNA was also found to be abundant in several other tumourigenic tissues, and in particular ovarian cancer, suggesting that PRK1 may govern similar molecular events in a wide range of cancers (Galgano *et al.*, 2009).

In contrast, migration, matrigel invasion and scratch wound assays showed that PRK2, and to a lesser extent PRK1, are vital for cell migration and invasion of a bladder tumour cell line and their roles in migration are also non-redundant (Lachmann *et al.*, 2011). However conflicting evidence showed that PRK2 may inhibit tumour growth in a mouse xenograft model (Cheng *et al.*, 2018). In this study, PRK2 in colon cancer cells was found to interact with and phosphorylate DUSP6 phosphatase, which inhibited the phosphorylation of Erk1/2. Erk1/2 were therefore no longer able to phosphorylate the CREB/Elk1 transcription factors thereby inhibiting the transcription of interleukins 4 and 10, which explains decreased levels of tumour-associated macrophages (TAMs) in human colon cancer. This anti-tumourigenic role of PRK2 may however be cell type-specific.

The role of PRK3 in cancer and specifically metastasis has seen increasing characterisation (Leenders *et al.* 2004). PRK3 was found to be necessary for the migration of PC-3 cells, a prostate cancer cell line, while PRK3 inhibition abrogated lymph node metastasis in an orthotopic mouse prostate tumour model. Contrastingly, O'Sullivan *et al.* (2015; 2017) found that PRK1 and PRK2, but not PRK3, were important for migration in PC-3 cells. The importance of PRK1 and PRK2 in prostate cancer cell migration has also been demonstrated by another recent study (Yang *et al.*, 2017).

Unsal-Kacmaz *et al.* (2012) confirmed the above findings of the role of PRK3 in primary tumour growth and lymph node metastasis in mice and interestingly identified RhoC as the preferred binding partner of PRK3 during metastasis. This is in contrast to PRK3 showing higher affinity for RhoB than RhoA or RhoC *in vitro* (Hutchinson *et al.*, 2013). The study also showed that PRK3 and RhoC are over-expressed in breast cancer cell lines and that PRK3 is critical for migration of these cells in a RhoC-dependent manner. This highlights that PRK3 is the key driver of neoplastic growth of both prostate and breast cancer cells. Additionally, PRK3 is vital for melanoma metastasis in mice and has been implicated in the glycosylation of the adhesion molecule ICAM-1, which is involved in signalling between tumour cells during metastasis (Mukai *et al.*, 2016). PRK3 therefore drives metastasis, a hallmark of cancer.

An attempt at targeting PRK1 has been made recently, involving the design of small molecule inhibitors which bind to the protein's kinase domain (Slynko *et al.*, 2016). Additionally, the low expression of PRK3 in most tissues, in contrast to its high expression in cancer, makes it an excellent therapeutic target. An attempt has been made to inactivate PRK3 using an siRNA/lipoplex complex known as Atu027, which showed efficacy in mouse prostate cancer and mouse pancreatic cancer models (Aleku *et al.*, 2008) and very low toxicity in patients (Strumberg *et al.*, 2012). This is currently in Phase II clinical trials.

1.4.7.2 Pathogens

Salmonella bacteria hijack PRK1 to activate an E3 ubiquitin ligase, SspH1. The crystal structure of the PRK1 HR1b domain in complex with SspH1 (Figure 1.15A) indicates that the HR1b domain interacts with the leucine-rich repeat (LRR) of this protein, which then

sterically displaces the catalytic domain bound to the LRR, causing its activation (Keszei *et al.*, 2014). Out of the HR1b residues involved in the interaction, 77% of them are in the domain's C-terminal helix, while only 4 or 5 of these residues are conserved in HR1a or HR1c, respectively, explaining the specificity of SspH1 for HR1b. The interacting residues are involved in a range of interactions including salt bridges, hydrogen bonds and hydrophobic interactions (Figure 1.15B).

The CagA toxin of *Helicobacter pylori*, known for disrupting cell adhesion and apical-basolateral cell polarity was found to co-immunoprecipitate with PRK2 (Mishra *et al.*, 2015). CagA was also responsible for the recruitment of PRK2 from the cytosol to the plasma membrane of AGS cells infected with *Helicobacter pylori*. Surprisingly, while CagA inhibited PRK2 activity, increased cell motility and cell invasive behaviour of cells infected with CagA were observed. This implicates PRK2 directly in the pathology resulting from this toxin. Other pathogens also depend on PRK2. These include the HCV virus, where PRK2 phosphorylates its RNA-dependent RNA polymerase so that it can replicate (Kim *et al.* 2012; Han *et al.*, 2014). Another example is the interaction of PRK2 with the leucine-rich repeat domains of the YopM proteins of *Yersinia pseudotuberculosis* and *Yersinia enterocolitica*; this interaction is necessary for full virulence (McPhee *et al.*, 2010; Hofling *et al.*, 2014). Whether this interaction involves binding to the HR1b domain has not been investigated.

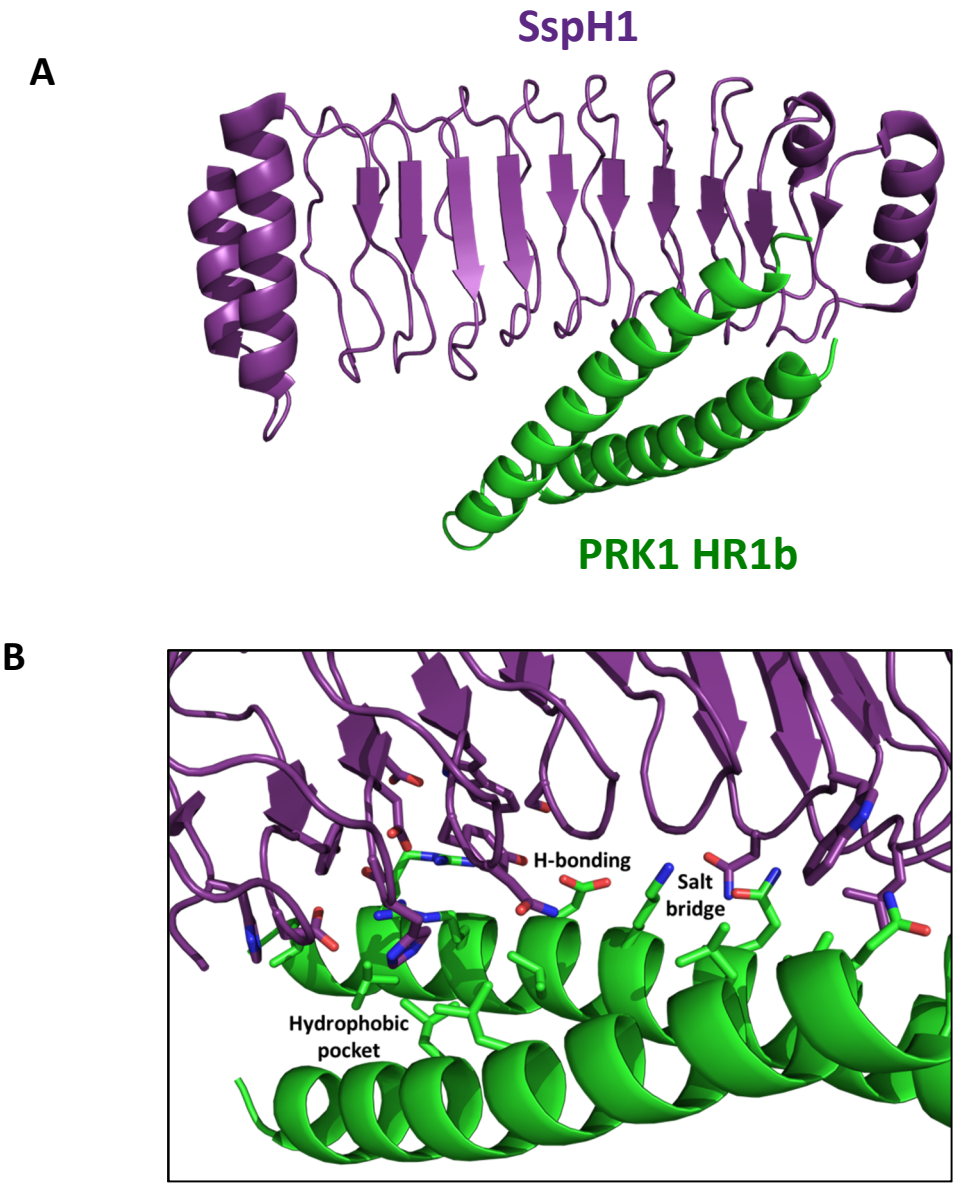


Figure 1.15 Structure of PRK1 HR1b in complex with the *Salmonella* protein SspH1. (**A**) Crystal structure of PRK1 HR1b (green) interacting with leucine-rich repeats of the *Salmonella* protein SspH1 (purple) (Keszei *et al.*, 2014) (PDB:4NKG). This interaction is important in the pathogenesis of *Salmonella*. (**B**) Key types of interactions are highlighted. Out of the 13 PRK1 HR1b residues involved in the interaction, 10 are found in the C-terminal helix.

1.4.8 HR1 domains as Rho family effector domains

1.4.8.1 Which HR1 domains bind to which Rho proteins

A feature of the PRKs is that they include three GTPase effector domains. This grants them the ability to potentially interact with a multitude of Rho GTPases. PRK1 was firstly identified as an effector of RhoA (Watanabe *et al.*, 1996). The HR1a domain was shown to bind to RhoA (Flynn *et al.*, 1998) and this was later determined to be an interaction of nM affinity (Owen *et al.*, 2003). The crystal structure of the RhoA/HR1a complex was solved by Maesaki *et al.* (1999), revealing two possible contact sites for RhoA, although only one of them is relevant in solution (Hutchinson *et al.*, 2011).

RhoA, RhoB and RhoC can all interact with the HR1a domain of PRK1, PRK2 and PRK3, indicating that PRKs are effectors of the entire Rho subfamily (Hutchinson *et al.*, 2013). These interactions also indicate that to some extent the interactions between HR1 domains and G proteins can be predicted based on conserved key residues. Differences between the affinities of these interactions do exist. PRK2 HR1a and PRK3 HR1a, as well as their respective HR1ab di-domains, bind to the three Rho GTPases with weaker affinities than their PRK1 counterparts. In addition, while the C-termini of the Rho proteins did not affect their affinity for PRK1 and PRK2 HR1 domains, the affinity of PRK3 HR1a and HR1ab for full length rather than truncated RhoB was higher. Also of interest was the affinity of RhoB for PRK1 HR1ab which was higher than that of the isolated HR1a domain. This was observed uniquely in the case of RhoB and not with RhoA or RhoC, suggesting possible cooperative binding of PRK1 HR1a and HR1b to RhoB. Significant differences in helicity and melting temperature were also noted between these HR1 domains, suggesting that distinct thermal stabilities, and as a result structural rigidity, may explain the accessibility of interacting residues by Rho GTPases and thus define the affinity of these interactions. Therefore Rho GTPase interactions with HR1 domains may involve use of conserved residues, the C-terminus of the GTPase, co-operative binding by two HR1 domains as well as protein stability and rigidity.

RhoA also binds to HR1b weakly (~6x more weakly than to HR1a (Owen *et al.*, 2003; Blumenstein & Ahmadian, 2004; Flynn *et al.*, 1998). Whether this interaction competes with

the RhoA/HR1a interaction or uses a different binding site has not been determined. The HR1ab di-domain binds to RhoA as tightly as HR1a (Owen *et al.*, 2003) indicating that there is no cooperative binding between the domains. The binding of RhoB and RhoC to HR1b alone has not been investigated.

Rac1 was found to interact with PRK2 (Vincent & Settleman, 1997). Subsequent work showed that Rac1 binds to both HR1a and HR1b (Owen *et al.*, 2003). The affinity of Rac1 for HR1b however is more than twofold higher than its affinity for HR1a. The affinity of Rac1 for the HR1ab di-domain is identical to the affinity for just the HR1b domain, suggesting that only one of the two HR1 domains may interact with a Rac1 molecule at the same time. The interaction of Rac1 with PRK3 has not been formally investigated, however the conservation of the key interacting residues between the HR1b domains indicates that PRK3 HR1b most probably does interact with Rac1 (Figure 1.16). The simultaneous binding of two Rac1 molecules to HR1a and HR1b or of RhoA and Rac1 to HR1a and HR1b, respectively, have not been investigated yet but remain a possibility.

No interacting partners are known for the HR1c domain of any of the PRK isoforms. HR1c may not interact with a GTPase or simply interacting partners have not been identified yet. The only study to investigate HR1c binding to GTPases came from Flynn *et al.* (1998) who demonstrated with a pull-down that RhoA does not bind to HR1c. The binding of HR1 domains to the rest of the Rho family proteins has not been investigated to date.

1.4.8.2 HR1 domain similarity and identity

The HR1 region is the Rho GTPase-interacting region of the PRKs. Each PRK has three HR1 domains, so there are a total of nine HR1 domains across all three PRKs. Each of these HR1 domains can potentially interact with one or more Rho GTPases.

Figure 1.16 shows a sequence alignment between the three HR1a, the three HR1b and three HR1c domains of PRK1, PRK2 and PRK3. Table 1.3 provides information on sequence identity and similarity between all nine HR1 domains. HR1 domains show ~60% sequence similarity and ~25% sequence identity. This divergence may explain their differential biophysical and GTPase-interacting properties (Hutchinson *et al.*, 2013).

PRK1	**HR1a**									
	HR1b	53 (24)								
	HR1c	38 (21)	53 (26)							
PRK2	**HR1a**	72 (45)	58 (23)	47 (19)						
	HR1b	55 (20)	81 (64)	59 (26)	56 (19)					
	HR1c	43 (25)	55 (26)	77 (55)	45 (18)	59 (24)				
PRK3	**HR1a**	69 (53)	45 (22)	49 (29)	74 (45)	55 (21)	54 (26)			
	HR1b	54 (29)	68 (44)	58 (29)	58 (28)	68 (43)	53 (26)	53 (28)		
	HR1c	49 (25)	51 (28)	72 (55)	50 (25)	56 (29)	68 (51)	46 (20)	56 (35)	
		HR1a	**HR1b**	**HR1c**	**HR1a**	**HR1b**	**HR1c**	**HR1a**	**HR1b**	**HR1c**
		PRK1			**PRK2**			**PRK3**		

Table 1.3 Sequence identity and similarity between the HR1 domains. The numbers at the top indicate the percentage sequence similarity (calculated as total of fully conserved or highly similar residues) while the number at the bottom in brackets indicates the percentage identity.

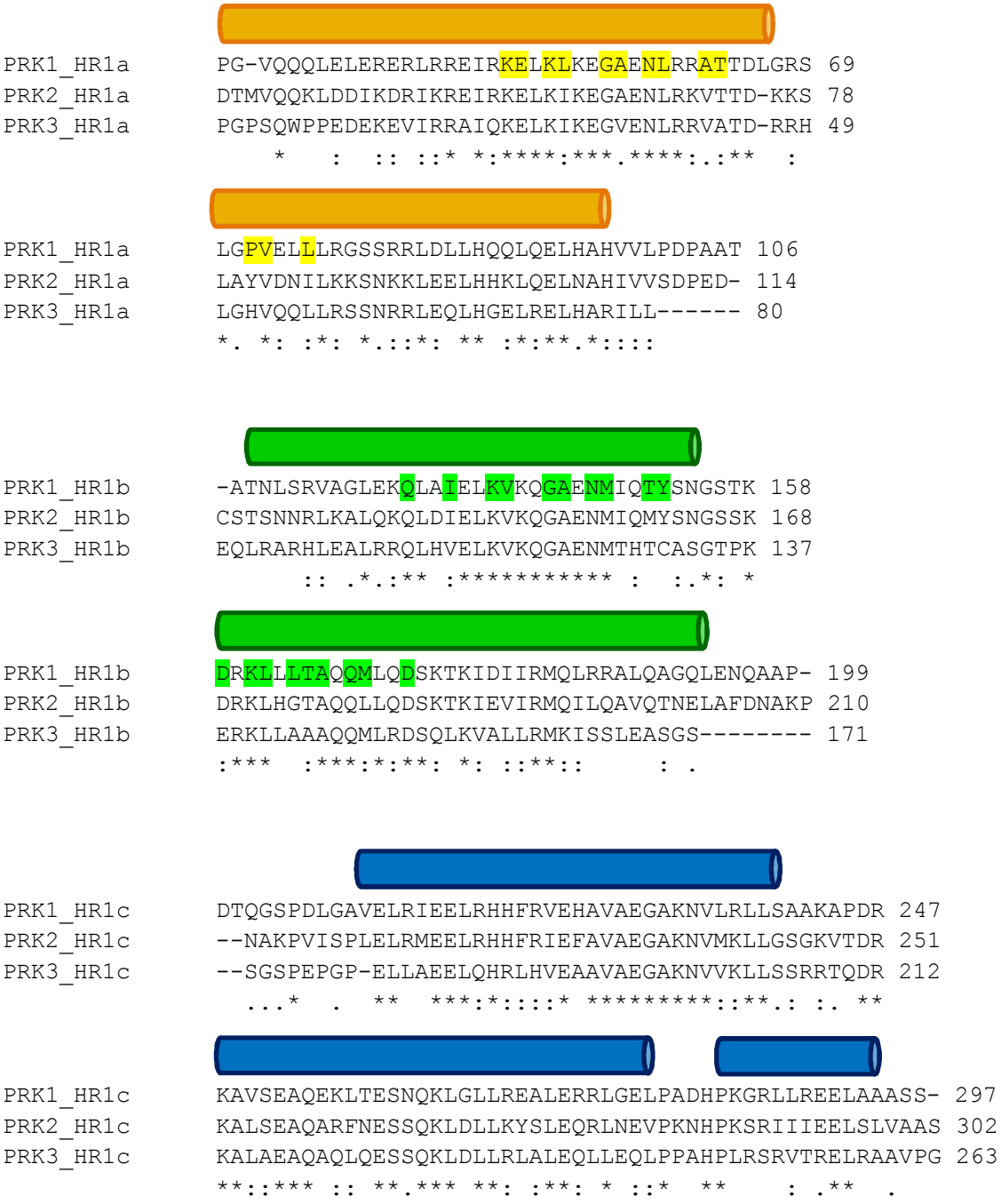


Figure 1.16 Sequence alignments of the HR1 domains of the three PRKs. There is conservation between the three HR1a domains, between the three HR1b domains and between the three HR1c domains. Asterisks highlight identical residues at all positions, colons highlight conservation of a strong group and periods highlight conservation of a weak group. The cylinders represent α-helices based on the published structures of PRK1 HR1a (Maesaki *et al.*, 1999) (PDB:1CXZ) and PRK1 HR1b (Modha *et al.*, 2008) (PDB:2RMK). No structure of an HR1c domain is currently available so the cylinders represent α-helices based on secondary structure prediction using jpred. Interacting residues are highlighted yellow for the RhoA/HR1a interaction and green for the Rac1/HR1b interaction.

1.4.8.3 Structural studies

The advent of the structural understanding of G protein/HR1 interactions came with the crystal structure of RhoA in complex with PRK1 HR1a (Maesaki *et al.*, 1999). HR1a was shown to form an antiparallel coiled coil structure, as predicted. Interestingly, HR1a was also shown to have a short N-terminal helix consisting of 5 amino acids which was not part of the interaction with RhoA. The structure identified two RhoA contact sites (I and II) for HR1a (Figure 17A). Contact site I buried a surface area of 2080 $Å^2$ and involved a series of salt bridges and an extensive hydrogen bond network between HR1a and RhoA α1, β2/β3 and α5 but no contacts with either of the switches. Contact site II buried a surface area of 1640 $Å^2$ and comprised mainly hydrophobic contacts between residues in the first long helix of HR1a and residues which were part of the switch regions in RhoA (Figures 1.17B and 1.17C). The authors believed that contact I was the physiological site based on the greater surface area that the interaction buried. However, Flynn *et al.* (1998) had shown that HR1a interacts with RhoA in a nucleotide-dependent manner, suggesting that contact II might be the true interaction site. Subsequent analysis of contact I and contact II RhoA mutants showed that contact I mutants did not affect HR1a binding, while some contact II mutants did (Hutchinson *et al.*, 2011). This indicated that contact site II is the most physiologically relevant. This should be unsurprising as it is the site which involves the switch regions and the interaction of RhoA with PRK1 has been shown to be nucleotide-dependent. This contact site also matches that used by other effectors such as ROCK when it binds to RhoA (Dvorsky *et al.*, 2004). A key question left unclear by this structural study is whether free HR1a forms a coiled coil in solution or is partially or fully unstructured and whether the N-terminal helix is an artefact of crystal packing in the complex structure.

The solution NMR structure of HR1b was solved and showed that HR1b also forms an antiparallel coiled coil (Owen *et al.*, 2003). Unlike HR1a, HR1b does not have an N-terminal helix. The solution NMR structure of Rac1 in complex with HR1b (Modha *et al.*, 2008) shows that HR1b interacts with Rac1 at a site which is equivalent to the contact II site of RhoA (Figure 1.18A), burying a surface area of ~1500 $Å^2$. This suggests that HR1 domains may all use this particular contact site of G proteins. Both helices of HR1b make extensive contacts with switch II and less so with switch I. The Rac1 switches make several contacts with HR1b around the mid-section and inter-helical loop of the domain and far away from the N- and

C-termini (Figures 1.18B and 1.18C). The interface mainly comprises hydrophobic interactions, e.g. Leu267^{Rac1} is buried in a hydrophobic pocket of HR1b which is formed by Val141^{HR1b} (helix 1), Ala145^{HR1b} (helix 1) and Met169^{HR1b} (helix 2). Hydrogen-bonding and salt bridge interactions are also employed but are less prevalent. The structure also highlights a key difference between the way Rac1 interacts with HR1b and the way RhoA interacts with HR1a. Rac1 has a triple-proline motif in its C-terminus which is followed by a polybasic region consisting of four arginines and two lysines. The prolines form a polyproline helix which forces the polybasic region into an extended conformation to avoid charge clashes. As a result, the polybasic region folds back and interacts with both the core G domain of Rac1 and with HR1b.

Currently, there are no reported structures of the PRK1 HR1c domain or of any of the PRK2 or PRK3 HR1 domains. The HR1c domain is predicted to have a unique C-terminal helix (Figure 1.16). It is unknown if this helix truly forms in solution and what structural and functional importance it may have. Furthermore, the structures of the entire HR1 region or of the full length PRK have not been investigated, thereby leaving possible inter-domain, intramolecular interactions a mystery. An additional unknown is highlighted by the higher affinity of RhoB for the HR1ab di-domain (Hutchinson *et al.*, 2013). This would suggest cooperative binding of RhoB by both HR1a and HR1b, although the structural basis of this finding has not been investigated.

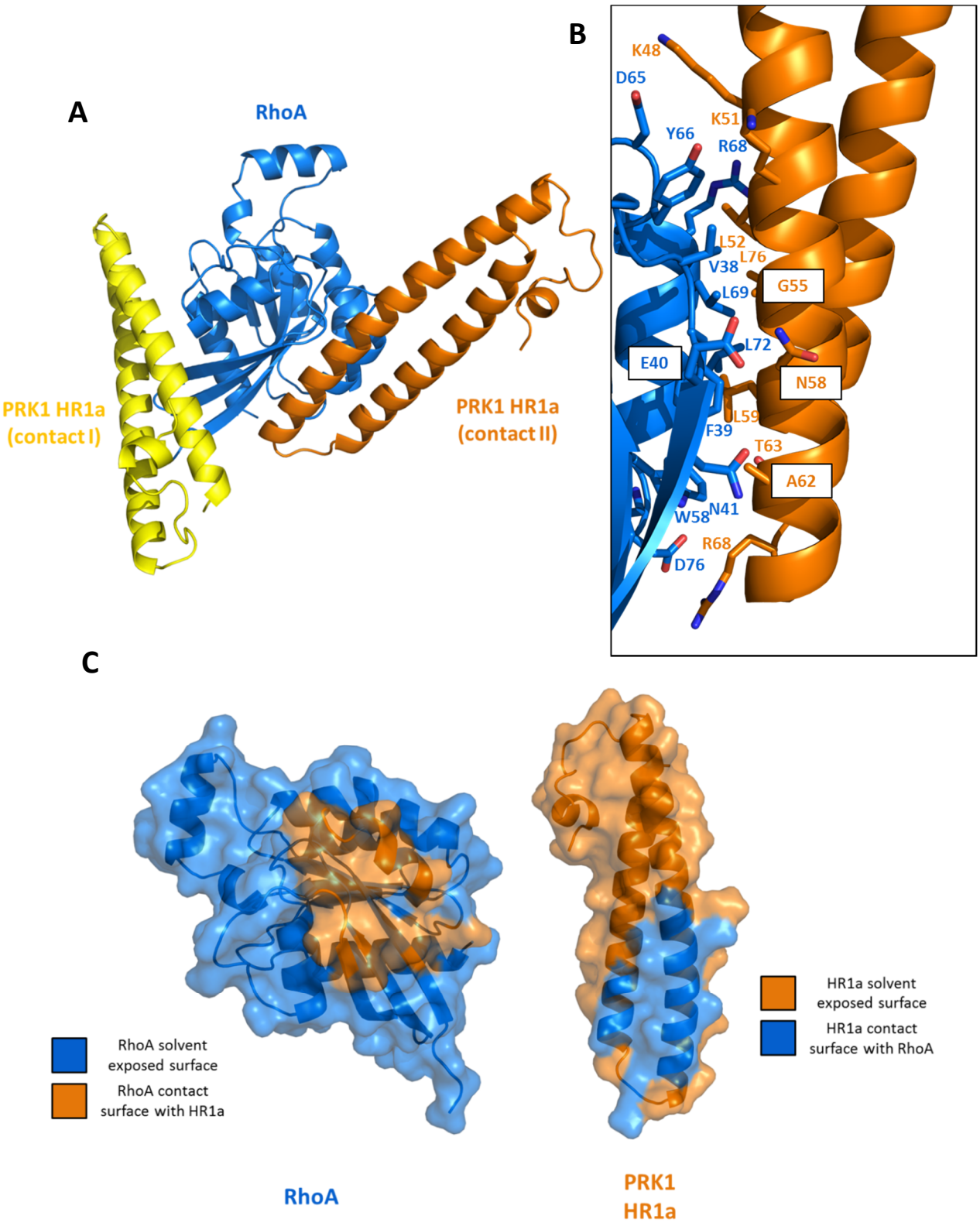


Figure 1.17 Structure of PRK1 HR1a interacting with RhoA. (**A**) Crystal structure of RhoA interacting with the HR1a domain of PRK1, which forms an antiparallel coiled-coil (Maesaki *et al.*, 1999) (PDB:1CXZ). Two different contacts, I and II, were identified in the crystals. (**B**) Most of the key residues in the contact II interface are highlighted. (**C**) Surface plots indicating the RhoA and HR1a surfaces used in the interaction.

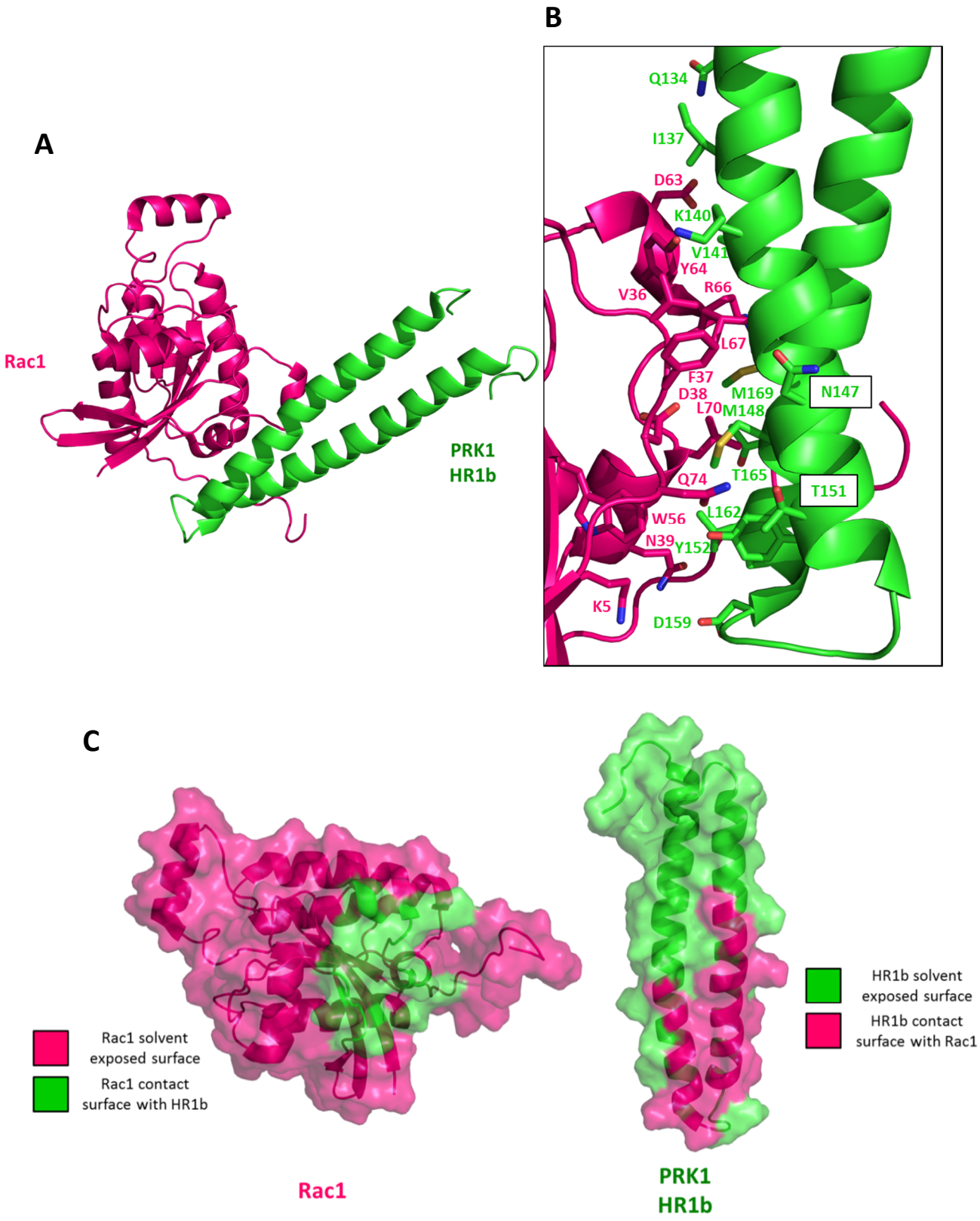


Figure 1.18 Structure of PRK1 HR1b interacting with Rac1. (**A**) Solution structure of Rac1 interacting with the HR1b domain of PRK1 (also an antiparallel coiled-coil) (Modha *et al.*, 2008) (PDB:2RMK). The contact interface between the two domains mirrors contact II of RhoA-HR1a (Figure 1.17A). (**B**) Most of the key residues in the contact interface are highlighted. (**C**) Surface plots indicating the Rac1 and HR1b surfaces used in the interaction.

1.5 PRKs - The uncertain and the unknown

1.5.1 The uncertain

1.5.1.1 Specificity and differential affinities of the G protein/HR1 domain interactions

While a range of Rho GTPases are known to interact with various HR1 domains, the specificity of these interactions remains undetermined and cannot be attributed to GTPase- and HR1-specific sequence motifs. This makes predicting novel GTPase/HR1 interactions a challenging task. In some cases the interaction seems likely, e.g. Rac1-interacting residues in PRK1 HR1b are conserved in PRK2 HR1b and PRK3 HR1b. The rigidity of coiled coils may also be a contributing factor to these interactions (Hutchinson *et al.*, 2013).

Surprisingly, it has been demonstrated that while RhoA interacts with HR1a and HR1ab with comparable affinities, RhoB interacts more tightly with the HR1ab di-domain (Hutchinson *et al.*, 2013). This suggests that both HR1a and HR1b could interact with RhoB at the same time. The most probable way this could happen is if both contact sites are used on RhoB, e.g. HR1a binds to contact II while HR1b binds to contact I. This cooperative binding has not been investigated yet but stands as another variable when one studies these interactions.

1.5.1.2 Activation of the PRKs

RhoA and PDK1 can both activate PRK1 via direct interactions. It remains uncertain as to whether there is a temporal order of these activation events or whether they can interact concomitantly. An additional undefined detail is whether RhoA is released upon PRK1 activation. Yoshinaga *et al.* (1999) identified three autophosphorylation sites in PRK1; one is a threonine in HR1a which makes a Van der Waals contact with a tryptophan residue in RhoA as indicated by the RhoA/HR1a structure. It is possible that activation of PRK1 leads to its autophosphorylation, with the phosphothreonine disrupting the interaction leading to RhoA dissociation.

Following activation at the cell membrane, the PRKs relocate to multiple cellular compartments, including the endosomes and the nucleus. It is possible that PRKs are guided

to the right compartment by a compartment-specific GTPase, e.g. RhoB localises PRK1 to the endosomes. Nuclear localisation may require a nuclear localisation signal if the PRKs do not enter the nucleus in complex with another protein, but no NLS signals have been identified yet.

1.5.1.3 Autoregulation of the PRKs

The pseudosubstrate region in the HR1a domain of PRK1 has been linked to the autoregulation of PRK1 catalytic activity (Kitagawa *et al.*, 1996). This infers an intramolecular interaction which keeps PRK1 in an autoinhibited state. No attempt has been made at solving the structure of the full length protein in this state and neither has it been investigated as to whether the HR1a domain undergoes a conformational change for this to occur, e.g. the coiled coil structure is disrupted to allow residues 39-53 to bind to the catalytic domain. Furthermore, PRK2 is autoregulated by inhibition resulting from oligomerisation (Bauer *et al.*, 2012). The dimerisation region is not part of a predicted structured domain but may be structured in solution, however this has not been investigated to date. Additional PRK oligomerisation sites, the effect of PTMs such as phosphorylation on oligomerisation, as well as heterodimerisation between PRK isoforms have not been investigated either.

1.5.2 The unknown

1.5.2.1 The HR1c domain

A lot about the PRKs currently remains unknown. Firstly, the role of the HR1c domain. No published studies provide structural characterisation of this domain from any of the three PRKs. Additionally, any GTPase interaction partners of this domain, if any, also remain unknown. The overall structure of the full-length PRK has not been investigated either, preventing any insight into intra-molecular interactions which may help understand the physiological role of this kinase. PRKs localise by their C2-like domain to the plasma membrane to be activated by RhoA and PDK1. Lipids such as arachidonic acid at the plasma

membrane also activate PRKs. It is possible that in the context of the membrane the PRK undergoes conformational changes which are essential for its activation. Overall, the lack of structural insight has been the Damocles' sword in understanding PRK (auto-)regulation.

1.5.2.2 The HR1abc tri-domain

A key question that is often overlooked is whether the three HR1 domains exist independently in solution or not. It is assumed that within HR1abc the three HR1 domains each form an antiparallel coiled coil. This may not be the case however as the HR1 domains could be forming dissimilar structures, e.g. the HR1 domains may form a single, very long coiled coil consisting of 3 HR1 helices interacting with the remaining 3.

1.5.2.3 Binding of G proteins to the HR1abc tri-domain

If the three PRK HR1 domains exist independently in solution, then it can be assumed that they can interact freely with their cognate G proteins. If the three HR1 domains interact with each other, it is possible that these intramolecular interactions hold the HR1abc tri-domain in a "closed conformation" which may open up, e.g. when RhoA binds. The intramolecular interactions may also compete with the G protein interactions, resulting in a lower affinity of the HR1abc tri-domain for G proteins than that of single HR1 domains. Another key question which remains unanswered is whether two G proteins can bind to HR1abc at the same time or whether this is sterically impossible. Structural insight into this will be key to unravelling the knot of Rho GTPase signalling and understanding convergence of GTPase signalling into a single effector, in this case the PRKs.

1.6 Aims and experimental techniques

1.6.1 Aims

This PhD project primarily aimed to provide a key missing piece of the PRK puzzle, which is that of the biophysical and structural understanding of the HR1 domains. Specifically, the project aimed to investigate the often over-looked HR1c domain, as well as provide structural insight of the three HR1 domains in the context of the HR1abc tri-domain. Additionally, the project aimed to discover novel G protein interacting partners for the HR1 domains and in particular for HR1c, and illuminate potential supercomplexes involving the PRKs and more than one G protein. The project's initial aims were as follows:

1) To biophysically characterise the HR1c domain as well as the HR1abc tri-domain

2) To solve the NMR solution structure of the HR1c domain and compare it to the structures of other HR1 domains and in particular see if the additional, predicted C-terminal helix truly forms in solution

3) To investigate interactions, if any, between the HR1 domains and, if the three HR1 domains are not independent, structurally characterise the HR1abc tri-domain

4) To investigate G protein binding to HR1c and to HR1abc, including obtaining a measure of the affinity of such interactions, as well as whether two G proteins can bind to the HR1abc tri-domain at the same time.

Chapter 3 answers question 1) by providing a comprehensive biophysical characterisation of PRK1 and PRK3 HR1 domains. The data obtained here provide part of the answer to question 3) and thus guide the experiments done in subsequent chapters. Chapter 4 addresses question 2) and, in addition to the solution structure of the HR1c domain, provides a useful tool to be used in the next chapter. Question 3) is addressed in Chapter 5,

where inter-domain interactions between the HR1 domains are investigated. This chapter also connects the two previous chapters together and establishes a coherent model. Chapter 6 then adds on to this model by investigating G protein binding to HR1 domains, thereby answering question 4) and providing a holistic understanding of HR1abc structure and function.

1.6.2 Experimental techniques

To address the aims outlined above, several biochemical and biophysical techniques were used. These techniques include chromatography, circular dichroism (CD), analytical ultracentrifugation (AUC), nuclear magnetic resonance (NMR), X-ray crystallography, isothermal titration calorimetry (ITC), chemical cross-linking, scintillation proximity assays (SPAs) and pull-downs. For each technique, sufficient information required to understand the results is succinctly provided in the appropriate chapter.

Chapter 2 Materials and methods

2.1 Materials

2.1.1 Chemicals

General chemicals were purchased from Sigma-Aldrich, Thermo Fischer, Formedium, Invitrogen, GE Healthcare and New England BioLabs. UltraPure™ Low Melting Point Agarose was purchased from Life Technologies. ^{15}N-NH_4Cl was purchased from Sigma-Aldrich and ^{13}C glucose was purchased from Cambridge Isotope Laboratories. BioRad reagent was purchased from Bio-Rad. Cross-linking agents EDC, DMS and DMP (Thermo Fischer) were a kind gift from Prof. Dame Jean Thomas.

2.1.2 Commercial kits and other materials

2.1.2.1 DNA purification

The GenElute Plasmid Miniprep kit was purchased from Sigma-Aldrich and used for plasmid purification from *E.coli*. The QIAquick Gel Extraction kit was purchased from QIAGEN and used for gel extraction of DNA from agarose gels.

2.1.2.2 TOPO cloning

The TOPO TA Cloning Kit containing pCR2.1 TOPO vector and One Shot TOP10F' chemically competent *E.coli* used in TOPO cloning were purchased from Invitrogen.

2.1.2.3 Site-directed mutagenesis

The QuikChange Lightning Multi Site-Directed Mutagenesis kit and ultracompetent XL10 Gold *E.coli* used for DNA mutagenesis were purchased from Agilent technologies.

2.1.2.4 PD10 columns

PD10 columns were purchased from GE Healthcare.

2.1.2.5 Bead slurries

Glutathione agarose used in protein affinity purification was purchased from Sigma-Aldrich. Alkaline phosphatase agarose was purchased from Sigma-Aldrich for use in nucleotide exchange reactions.

2.1.3 Commercial enzymes

Restriction enzymes (*Bam*HI and *Eco*RI) and T4 DNA ligase were purchased from Roche. Calf-intestinal alkaline phosphatase was purchased from Sigma-Aldrich and used in the preparation of cut vector. Thrombin was purchased from Merck Millipore.

2.1.4 Oligonucleotides

Oligonucleotides for PCR were purchased from Sigma-Aldrich (see section 2.1.9).

2.1.5 Microbiological media

2TY	1.6% Tryptone, 1% Yeast extract, 0.5% NaCl
LB-agar	1% Tryptone, 0.5% Yeast extract, 1% NaCl, 1.6% Bacto-agar
Antibiotics (working concentrations)	Ampicillin: 100 µg/ml Chloramphenicol: 20 µg/ml Kanamycin: 25 µg/ml Tetracycline: 10 µg/ml
MOPS	10x MOPS: 400 mM K-MOPS pH 7.4, 40 mM Tricine pH 7.4, 5.28 mM $MgCl_2$, 2.76 mM K_2SO_4, 500 mM NaCl, 100 µM $FeSO_4$, 5 µM $CaCl_2$ 100x micronutrients: 0.4 mM H_3BO_3, 80 µM $MnCl_2$, 30 µM $CoCl_2$, 10 µM $CuSO_4$, 10 µM $ZnSO_4$, 3 µM $NH_4(MoO_2)_2$ MOPS (per litre): 100 ml 10x MOPS, 1 ml 100x micronutrients, 1.32 mM KH_2PO_4 pH 7.4, 0.4% (w/v) glucose, 20 mM NH_4Cl, 5% (v/v) Celtone

	(Cambridge Isotopes Ltd)
M9	47 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 4 µM $ZnSO_4$, 1 µM $MnSO_4$, 0.7 µM H_3BO_4, 0.7 µM $CuSO_4$, 0.4% (w/v) glucose, 2 µM $FeCl_3$, 2 µM $MgSO_4$, 20 mM NH_4Cl, 0.1 µM $CaCl_2$

Table 2.1 Microbiological media.

2.1.6 Buffers

2.1.6.1 Cloning and transformation buffers

Agarose gel buffers	
50x TAE buffer	2 M Tris base, 5% acetic acid, 50 mM EDTA pH 8.0
5x STOP buffer	50% sucrose (w/v), 50 mM EDTA pH 7.5, 0.1% bromophenol blue
Transformation buffers	
TF1	10 mM $CaCl_2$, 10 mM MOPS-KOH, pH 6.8
TF2	75 mM $CaCl_2$, 10 mM MOPS-KOH, pH 6.8
TF3	75 mM $CaCl_2$, 5% glycerol, 1 mM EDTA, 50 mM MOPS-KOH, pH 7.5

Table 2.2 Cloning and transformation buffers.

2.1.6.2 Protein purification buffers

Buffer	**Components**
MTPBS	20 mM sodium phosphate pH 7.3, 150 mM NaCl
MTPBS (High salt)	20 mM sodium phosphate pH 7.3, 500 mM NaCl
Thrombin cleavage buffer	20 mM Tris-HCl pH 8.4, 150 mM NaCl, 2.5 mM $CaCl_2$
Buffer A1	50 mM Tris-HCl pH 7.5, 0.05% NaN_3
Buffer A2	50 mM Tris-HCl pH 7.5, 5 mM DTT, 0.05% NaN_3
Buffer B1	50 mM Tris-HCl pH 7.5, 2 M NaCl, 0.05% NaN_3
Buffer B2	50 mM Tris-HCl pH 7.5, 2 M NaCl, 5 mM DTT, 0.05% NaN_3
Buffer S1	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% NaN_3
Buffer S2	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT, 0.05% NaN_3
Buffer S3	50 mM Tris-HCl pH 7.5, 500 mM NaCl, 5mM

	DTT, 0.05% NaN_3
Buffer G	50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM reduced glutathione, 0.05% NaN_3
TNMD	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT
$TNMD_{250}$	50 mM Tris-HCl pH 7.5, 250 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT
Buffer S4	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 0.05% NaN_3

Table 2.3 Protein purification buffers.

2.1.6.3 Biochemical and biophysical analysis buffers

Buffer	Components
CD buffers	
CD buffer 1	50 mM sodium phosphate pH 7.5, 150 mM sodium fluoride
CD buffer 2	50 mM sodium phosphate pH 7.5, 150 mM sodium fluoride, 1 mM DTT
CD buffer 3	50 mM sodium phosphate pH 7.5, 500 mM sodium fluoride, 1 mM DTT
Nucleotide exchange and HPLC buffers	
Nucleotide exchange reaction mix	5.29 mM GMPPNP, 217 µM $(NH_4)_2SO_4$, 20.6 mM Tris-HCl pH 7.5, 59.7 mM NaCl, 1.99 $MgCl_2$, 1.99 mM DTT
HPLC buffer	0.6 M ammonium phosphate pH 4.0
SPA exchange buffer	50 mM Tris-HCl pH 7.5, 150 mM NaCl
SPA buffers	
SPA reaction mix	5 mg/ml Protein A SPA fluoromicrospheres (GE Healthcare), 30 mM Tris-HCl pH 7.5, 0.17 mg/ml BSA, 0.9 mM $MgCl_2$, 1.7 mM DTT, 18.7 µg/ml anti-GST antibody (Life Technologies A5800) or 10 µg/ml anti-His (Sigma-Aldrich H1029)
SPA dilution buffer	50 mM Tris-HCl pH 7.5, 0.2 mg/mL BSA, 2 mM DTT, 1 mM $MgCl_2$
ITC buffers	
ITC buffer 1	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT
ITC buffer 2	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT
Cross-linking buffers	
EDC buffer	100 mM MES pH 5.0, 150 mM NaCl, 5 mM DTT
DM buffer	100 mM sodium phosphate pH 8.0, 150 mM NaCl, 5 mM DTT

Table 2.4 Biochemical and biophysical analysis buffers.

For all other biochemical and biophysical analyses, the buffers used were identical to the final purification buffer of each protein, unless stated otherwise.

2.1.6.4 SDS-PAGE buffers

Laemmli gels for sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)		
Resolving gel	**Component**	**Volume (ml)**
	1.5 mM Tris-HCl pH 8.8	1.5
	Acrylamide:Bis-Acrylamide (29:1, 30% (w/v))	2.4-3.6 (12-18%)
	MilliQ H_2O	0.8-2.0
	10% (w/v) SDS	0.06
	10% (w/v) APS	0.048
	TEMED	0.006
	Total	6
Stacking gel	**Component**	**Volume (ml)**
	0.5 mM Tris-HCl pH 6.8	0.5
	Acrylamide:Bis-Acrylamide (29:1, 30% (w/v))	0.25
	MilliQ H_2O	1.2
	10% (w/v) SDS	0.02
	10% (w/v) APS	0.04
	TEMED	0.002
	Total	2
2x SDS Loading buffer	3.3% (w/v) SDS, 6 M Urea, 33.3 mM Tris-HCl pH 7.4, 66.7 mM β-mercaptoethanol, 0.1% bromophenol blue	
Running buffer	25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS	

Table 2.5 SDS-PAGE buffers.

2.1.7 Plasmids

2.1.7.1 Vectors

Vector	**Fusion tag**	**Protease cleavage site**	**Source**	**Antibiotic resistance**	**Bacterial cell type**	**[IPTG] (mM)**
pCR 2.1 TOPO	-	-	Invitrogen	Ampicillin/ Kanamycin	TOP10 F'	N/A
pGEX-6P-1	GST (N)	3C protease	GE Healthcare	Ampicillin	XL1, XL10 Gold, BL21, Rosetta2, C41	0.1

pGEX-HISP	GST (N), His_6 (C)	3C protease	Based on pGEX-6P-1 (Hutchinson *et al.*, 2011)	Ampicillin	XL1, XL10 Gold, BL21, Rosetta2	0.1
pGEX-2T	GST (N)	Thrombin	GE Healthcare	Ampicillin	BL21	0.1

Table 2.6 DNA plasmid vectors. N stands for N-terminal and C for C-terminal.

2.1.7.2 Constructs

The generated constructs are listed in Table 2.7. Expression vectors already available in the lab are in italic. Other constructs used as templates in PCR or for mutagenesis (e.g. pGEXHISP constructs) were already available in the lab from the Part III Master's Project (Sophocleous, 2014) or had been cloned previously by Catherine Hutchinson, so they are not listed here. All sequences are from *Homo sapiens*. Expression vectors were always cloned with *Bam*HI and *Eco*RI restriction sites.

Vector	Construct	Construct limits
pCR 2.1 TOPO	PRK1 HR1abc	1-297
	PRK1 HR1bc	122-297
	RhoA Δ4 F25N Q63L	1-194
	RhoD Δ4 Q75L	1-211
pGEX-6P-1	*PRK1 HR1a*	1-106
	PRK1 HR1ab	1-199
	RhoA Δ4 F25N Q63L	1-194
	RhoD Δ4 Q75L	1-211
pGEX-HISP(STOP)	PRK1 HR1b	122-199
	PRK1 HR1c	201-297
	PRK1 HR1bc	122-297
	PRK1 HR1abc	1-297
	PRK2 HR1c	207-302
	PRK2 HR1bc	132-302
	PRK2 HR1abc	1-302
	PRK3 HR1a	1-80
	PRK3 HR1b	100-171
	PRK3 HR1c	169-263
	PRK3 HR1ab	1-171

	PRK3 HR1bc	100-263
	PRK3 HR1abc	1-263
pGEX-2T	*RhoA Δ12 F25N Q63L*	1-186
	Cdc42 Δ7 Q61L	1-190

Table 2.7 Table of constructs.

2.1.7.3 Sequences

The amino acid sequences of all purified proteins, as well as vector tags are listed in Appendix 1.

2.1.8 Bacterial strains

Table 2.8 lists the *E.coli* strains used, their antibiotic resistance and their genotype.

Strain	Supplier	Antibiotic resistance	Genotype	Phenotype
XL1 MRF^{+} Blue	Stratagene	Tetracycline	*endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [*F´ *proAB lacI*qZΔ*M15* Tn*10* (Tet^{R})	Deficient in DNA recombinase, DNA endonuclease and DNA gyrase and also contains *lac*/Z mini-gene for blue-white selection.
TOP10 F’	Invitrogen	Tetracycline	F’[*lac*I Tn*10* (Tet^{R})] *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139	Deficient in DNA recombinase. The F’ episome contains the *lac*I^{q} repressor for inducible expression with IPTG and blue-white screening.
XL10 Gold	Agilent technologies	Tetracycline	TetrΔ(*mcr*A)183 Δ(*mcr*CB-*hsd*SMR-*mrr*)173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac*Hte [F’ *pro*AB *lacI*qZΔM15 Tn*10* (Tet^{R})]	Very high transformation efficiency. Deficient in DNA endonuclease, DNA recombinase and DNA gyrase.
BL21	Invitrogen	None	F^{-} *ompT hsdS*(r_{B}^{-} m_{B}^{-})	Deficient in ompT and proteases

BL21-Rosetta2 (DE3) pLysS	Novagen	Chloramphenicol	F$^-$ *ompT hsdS*(r_B^- m_B^-) *gal dcm* (DE3) plysSRARE2 (Cam^R)	Based on BL21. Contains DE3 lysogen which encodes T7 polymerase and also contains pRARE2 plasmid from which rare tRNAs for 5 amino acids (Gly, Pro, Arg, Ile, Leu) are constitutively expressed.
BL21 (DE3) C41	Sigma-Aldrich	None	F$^-$ *ompT hsdS*(r_B^- m_B^-) *gal dcm* (DE3)	Contains uncharacterised mutations which help with the expression and solubility of "toxic" proteins.

Table 2.8 Bacterial strains.

2.1.9 Primers

Conventional cloning and mutagenic primers are listed in Table 2.9.

Conventional cloning primers	
Primer	**Sequence (5'-3')**
PRK1 HR1abc F'	CGCGGATCCCATATGGCCAGCGACGCC
PRK1 HR1abc R'	GCGGAATTCGGGGAGGAGGCCGCAGC
PRK1 HR1bc F'	CGCGGATCCCATATGGCCACCAACCTGAGCC
PRK1 HR1bc R'	GCGGAATTCGGGGAGGAGGCCGCAGC
RhoA F'	CGCGGATCCATGGCTGCCATCCGG
RhoA R'	CGCGAATTCTCATCACCCAGATTTTTTCTTCCC
RhoD F'	CGCGGATCCATGACGGCGGCCCAG
RhoD R'	CGCGAATTCTCATCAAAAGCCCTGGGTAATCCG
Mutagenic primers	
Primer	**Sequence (5'-3')**
PRK1 HR1abc stop	GCTGCGGCCTCCTCC**TGA**AATTCCCGGGTCGACTCG
PRK1 HR1b stop	GCGCTGCAGGCCGGCCAGCTGGAGAACCAGGCAGCCCCG**TGATGA**TCGCCCTTCGCG

PRK1 HR1c stop	GCTGCGGCCTCCTCC**TGA**AATTCCCGGGTCGACTCG
PRK1 HR1bc stop	GCTGCGGCCTCCTCC**TGA**AATTCCCGGGTCGACTCG
PRK2 HR1c stop	CACTTGTTGCTGCATCA**TGA**AATTCCCGGGTCGACTCG
PRK2 HR1bc stop	CACTTGTTGCTGCATCA**TGA**AATTCCCGGGTCGACTCG
PRK2 HR1abc stop	CACTTGTTGCTGCATCA**TGA**AATTCCCGGGTCGACTCG
PRK3 HR1a stop	CACGCCCGAATCCTGCTG**TGATGA**TCCCGGGTCGACTCGAGC
PRK3 HR1b stop	CTGGAGGCCAGTGGGTCC**TGATGA**TCCCGGGTCGACTCGAGC
PRK3 HR1c stop	GCTGCGGTGCCTGGA**TGA**AATTCCCGGGTCGACTCG
PRK3 HR1ab stop	CTGGAGGCCAGTGGGTCC**TGATGA**TCCCGGGTCGACTCGAGC
PRK3 HR1bc stop	GCTGCGGTGCCTGGA**TGA**AATTCCCGGGTCGACTCG
PRK3 HR1abc stop	GCTGCGGTGCCTGGA**TGA**AATTCCCGGGTCGACTCG

Table 2.9 Table of DNA primers. *Bam*HI restriction sites are highlighted in cyan, *Eco*RI restriction sites are highlighted in green, 5' CGC overhangs are highlighted in yellow and stop codons are in black.

2.2 Methods

2.2.1 Molecular cloning

2.2.1.1 Agarose gel electrophoresis

Agarose gels (0.7-1.5%) were made in TAE buffer with 0.5 µg/ml ethidium bromide. DNA samples were made up in 5x STOP buffer and 10-15 µl were loaded onto the gel. The gels were run in TAE buffer with 0.5 µg/ml ethidium bromide at 100 V, 100 mA for 1 h 30 min and were then visualised under UV light.

2.2.1.2 Polymerase Chain Reaction (PCR)

For each reaction, 200 µM dNTPs, 10 µl 10xTAQ Buffer (Roche), 100 pmol forward primer, 100 pmol reverse primer (Table 5) and 50 ng template DNA were combined and the volume adjusted to 99 µl with sterile AnalaR^{TM} water. 1 µl Taq Polymerase (Roche) was added after step 2 of the PCR program (Table 2.10). PCR products were visualised on a 1.5% agarose gel (100 V, 80 min) and verified according to size.

Step	Process	Temperature (^{o}C)	Time (min)	Cycles
1	Denature	95	6	1
2	Anneal	55	2	1
3	Pause and add 1 µl Taq Polymerase	72	-	-
4	Extend	72	2	25
5	Denature	95	2	
6	Anneal	55	2	
7	Extend	72	10	1
8	HOLD	4	HOLD	-

Table 2.10 PCR cycling program.

2.2.1.3 TOPO cloning

The PCR product was cloned into a pCR2.1-TOPO vector using a TOPO Cloning Kit (Life Technologies) and the resulting product was used to transform TOP10F' cells according to the manufacturer's instructions. The cells were plated on agar containing kanamycin, tetracycline, 500 µM IPTG (isopropyl β-D-1-thiogalactopyranoside) and 80 µg/ml X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). White colonies, suggesting successful incorporation of PCR product into the pCR2.1-TOPO vector, were then patched onto new agar plates.

2.2.1.4 Restriction digest of plasmid DNA

Overnight cultures (5 ml) were inoculated from white colonies. Bacterial plasmid DNA was purified using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich) according to the manufacturer's instructions.

DNA concentration was estimated by UV absorbance at 260 nm and calculated by

$$C = A_{260} \times 50\ \mu g/ml$$

where C is the concentration of DNA in µg/ml.

1 µl *Bam*HI (10 units) and 1 µl *Eco*RI (10 units) restriction enzymes and 2 µl Buffer B (Roche) were added to 16 µl DNA. Restriction digests were carried out for 2-4 hours at 37 ^{o}C and quenched by addition of 5 µl 5x STOP buffer. "Uncut" samples were prepared by adding 2.5 µl 5x STOP buffer to 10 µl DNA. Restriction digest products were visualised on a 1.5% agarose gel (100 V, 80 min) to confirm the presence of the correct insert in the vector and the insert sequence was confirmed by sequencing (Sequencing Facility, Department of Biochemistry).

2.2.1.5 Competent cells

10 ml 2TY with the appropriate antibiotic was inoculated with the relevant *E.coli* and incubated at 37 ^{o}C overnight shaking at 180 rpm. 0.75 ml of overnight culture was added to 20 ml 2TY and shaken at 37 ^{o}C to an A_{600} of 0.4-0.6, spun at 8300 x g for 10 min and the supernatant discarded. The pellet was resuspended in 10 ml of ice cold TF1 and spun at 7000 x g for 10 min. The supernatant was discarded and the pellet was resuspended in 1 ml of ice cold TF2 and left on ice for 3-24 h.

2.2.1.6 Bacterial transformation

Transformation of XL1-Blue MRF^{+} (hereafter XL1) competent cells was carried out by mixing 200 µl cells, 20 µl of the ligation reaction and 180 µl TF3. Transformation of BL21, C41 and Rosetta2(DE3) competent cells was carried out by mixing 100 µl cells, 1 µl DNA and 99 µl TF3 buffer. The cells were incubated on ice for 45 min, followed by heat shock at 42 ^{o}C for 10 min and a further incubation on ice for 2 min. 1 ml 2TY was then added and the cells were incubated at 37 ^{o}C for 30 min. 200 µl XL1 cells or 100 µl BL21, C41 or Rosetta2(DE3) cells were then spread on plates containing the appropriate antibiotic. The rest of the cells were pelleted at 16060 x g for 5 min and the supernatant discarded. The pellet was

resuspended in 200 µl 2TY and spread on a separate plate containing the appropriate antibiotic. Table 2.6 indicates which bacterial strain was used for which vector.

2.2.1.7 Preparation of linearised expression vector by alkaline phosphatase treatment

Plasmid DNA (pGEXHISP or pGEX6P) from 5 ml of overnight culture was purified and 5 µg were mixed with 10 µl Buffer B, 5 µl *Bam*HI (50 units) and 5 µl *Eco*RI (50 units) (Roche) and the volume adjusted to 100 µl with sterile AnalaR^{TM} water. The restriction digest was incubated for 4-6 h at 37 ^{o}C. Uncut and cut vector samples were visualised on a 0.7% agarose gel (100 V, 80 min) to confirm that the vectors had been digested. 71.4 µM Tris-HCl pH 8.0, 7.14 µM $MgCl_2$ and 1 unit calf-intestinal alkaline phosphatase were combined and the volume adjusted to 210 µl with sterile AnalaR^{TM} water. The reaction mixture was incubated at 37 ^{o}C for 20 min followed by warming to 65 ^{o}C for 20 min. The volume was adjusted to 360 µl with sterile AnalaR^{TM} water. 300 µl phenol was added to the reaction mixture which was vortexed for 30 sec. The mixture was spun at 16060 x g for 1 min and the aqueous layer containing the DNA was removed. The process was repeated by adding 300 µl phenol to the DNA which was then separated from phenol. The volume of the reaction was then adjusted to 610 µl by adding 300 µl chloroform and the mixture was vortexed for 30 sec. The mixture was spun at 16060 x g for 1 min and the aqueous layer containing the DNA was removed. The process was repeated by adding 300 µl chloroform to the DNA which was then separated from chloroform. 300 µl ether were then added to the aqueous layer and the mixture was vortexed for 30 sec. The mixture was spun at 16060 x g for 1 min and the top ether layer was removed. Residual ether was allowed to evaporate in a fume cupboard for 30 min. DNA was quantified as described in section 2.2.1.4.

2.2.1.8 Ligations

Plasmid DNA from 20 ml of overnight culture was purified and 5 µg were mixed with 10 µl Buffer B, 5 µl BamHI (50 units) and 5 µl EcoRI (50 units) (Roche) and the volume adjusted to 100 µl with sterile AnalaR^{TM} water. The restriction digest was incubated for 4-6 h at 37 ^{o}C and quenched by addition of 25 µl 5x STOP buffer. The products were visualised on a 1.5% UltraPure^{TM} Low Melting Point Agarose (Life Technologies) gel (100 V, 80 min) and the insert

fragment purified using the QIAquick Gel Extraction Kit (QIAGEN) according to the manufacturer's instructions.

16 µl insert, 1 µl alkaline phosphatase-treated, linearised expression vector (pGEXHISP or pGEX6P), 2 µl 10xT4 ligase buffer and 1 µl T4 ligase (Roche) were incubated at 16 ^{o}C for 16 h. XL1 competent cells were transformed with the ligation products and spread on agar plates containing ampicillin and tetracycline. Successful ligation reactions were identified by restriction digest of plasmid DNA from transformed XL1 colonies and confirmation of an insert of the correct size in the expression vector (as described in section 2.2.1.4).

2.2.1.9 Sequencing

DNA sequencing was performed by the DNA sequencing facility at the Department of Biochemistry, University of Cambridge. The sequencing files were viewed with Chromas Lite v2.1.1.

2.2.2 Site-directed mutagenesis

Site-directed mutagenesis was performed using the Agilent QuikChange Kit according to the manufacturer's instructions and using the mutagenic primers listed in Table 2.9 (see section 2.1.9).

2.2.3 SDS-PAGE

Laemmli gels of the appropriate percentage were poured in 1 mm gel cassettes (Thermo Fischer) according to Table 2.5. Gel samples were prepared in 2x SDS Loading Buffer, boiled for 5 min, spun at 16060 x g for 5 min and 10-15 µl loaded onto the gel alongside unstained broad range protein markers (New England Biolabs; markers used are indicated on each gel separately). Gels were run at 200 V, 45-55 mA (per gel) for 1 h 10 min, stained with InstantBlue (Expedeon) overnight and washed in water overnight.

2.2.4 Expression trials

10 ml 2TY with the appropriate antibiotic(s) was inoculated from a plate and incubated at 37 ^{o}C overnight shaking at 180 rpm. 10 ml 2TY with the appropriate antibiotic(s) was inoculated with 1 ml of overnight culture and shaken at 180 rpm, 37 ^{o}C. Three such cultures were grown to an A_{600} of 0.8. Two of the cultures were induced by adding 0.1 mM IPTG. One of the induced cultures was shaken at 180 rpm, 37 ^{o}C for 5 h. The uninduced and other induced cultures were shaken at 180 rpm, 20 ^{o}C for 16 h.

The total resuspension volume of MilliQ H_2O and 2x SDS Loading Buffer used was normalised to A_{600} x $\frac{1000}{6}$ µl. Uninduced and induced samples were prepared. 1 ml of each culture was centrifuged at 16060 x g for 5 min, the supernatant discarded and the pellets resuspended in half a volume MilliQ H_2O. Half a volume 2x SDS Loading Buffer was then added. Samples were also prepared by pelleting 2 ml of *E.coli* culture, discarding the supernatant and resuspending in a total volume MilliQ H_2O. 1 µl DNase (at 10 mg/ml) and 2 µl lysozyme (at 20 mg/ml) were added and the samples were left standing at room temperature for 20 min. The samples were sonicated (2x 30 sec, with 30 sec rest between each pulse) with 50% amplitude on a Sonic Dismembrator Model 120 (Thermo Fischer) and then centrifuged at 16060 x g for 5 min. 100 µl of the supernatant were transferred to a separate tube and 100 µl 2x SDS Loading Buffer was added to produce a soluble sample. The remaining supernatant was discarded and the pellet resuspended in a total volume MilliQ H_2O and a total volume 2x SDS Loading Buffer. All samples were boiled at 100 ^{o}C for 5 min, centrifuged at 16060 x g for 5 min and 15 µL loaded onto a gel. Gels were run at 200 V for 70 min.

2.2.5 Protein expression

8 x or 12 x 50 ml of 2TY with the appropriate antibiotic(s) was inoculated from a plate and incubated at 37 ^{o}C overnight shaking at 180 rpm. 8 x or 12 x 500 ml 2TY with the appropriate antibiotic(s) was inoculated with the overnight cultures. The bacterial cultures were grown

at 37 ^{o}C to an A_{600} of 0.8, cooled in cold water baths and induced by adding 0.1 mM IPTG. The cultures were then incubated at 20 ^{o}C for 20 h shaking at 180 rpm.

Bacterial cultures were centrifuged at 8671 x g for 15 min. Cell pellets were re-suspended in ice cold MTPBS (High salt) or $TNMD_{250}$ (in which a SIGMAFAST^{TM} Protease Inhibitor Cocktail Tablet, EDTA-Free (Sigma-Aldrich) had previously been dissolved), and 1 mM PMSF was then added. The cells were lysed by passing five times through an Avestin EmulsiFlex-C5 Emulsifier (ATA Scientific).

An additional 1 mM PMSF was added to the cell lysate which was then centrifuged at 40000 x g for 30 min. The supernatant was added to 2 x (for 4 L preps) or 3 x (for 6 L preps) 7.5 ml of glutathione agarose bead slurry (Sigma-Aldrich), which had previously been equilibrated in MTPBS (High salt) or $TNMD_{250}$ and was then left rotating at 4 ^{o}C for 2 h.

2.2.6 Protein purification

2.2.6.1 Glutathione agarose purification and proteolytic removal of N-terminal tag

The bead slurry was spun at 1000 x g for 10 min, the supernatant removed and a gel sample taken (10 µl of 1 in 2 dilution in 2x SDS Loading Buffer). The beads were then washed 3 times in MTPBS (High salt) or $TNMD_{250}$ and finally resuspended in an equal volume of MTPBS (High Salt) or $TNMD_{250}$. A gel sample of each wash supernatant was taken (10 µl of 1 in 2 dilution in 2x SDS Loading Buffer) and 10 µl were loaded on a gel. Gels were run at 200 V for 70 min. Additional MTPBS (High salt) buffer was then added up to 20 ml, followed by the addition of 1 mM DTT and 1 mM EDTA pH 7.4 (not for G proteins). 100 µl of HRV 3C protease at 4.8 mg/ml were then added and the tubes were incubated on a rotating wheel at 4 ^{o}C for 16 h. In the case of proteins with a thrombin cleavage site rather than a 3C protease cleavage site, the beads were washed 3x in thrombin cleavage buffer as above and 50 u of thrombin were added to each tube.

2.2.6.2 Size-exclusion chromatography

Size exclusion columns were equilibrated with 2.5 column volumes (CV) of the appropriate purification buffer on a Fast Protein Liquid Chromatography (FPLC) ÄKTA Purifier system (GE Healthcare). The sample was manually loaded using a 2 ml loop, which had previously been washed with 6 ml of purification buffer. The sample was injected on the column and eluted in 1.5 CV at a flow rate of 1 ml/min. 2 ml fractions were collected and subsequently analysed by SDS-PAGE. The size exclusion columns used are listed in Table 2.11.

Column	Column volume (ml)	Manufacturer	Flow rate (ml/min)	Fraction volume (ml)
Superdex75 16/60	120	GE Healthcare	1	2
Superdex200 16/60	120	GE Healthcare	1	2
Heparin	20	GE Healthcare	1	1
ResourceQ	6	GE Healthcare	1	1
Glutathione sepharose	10	GE Healthcare	2	-

Table 2.11 Chromatography columns.

2.2.6.3 Ion exchange chromatography

Cation exchange chromatography was performed on a Heparin column while anion exchange chromatography was performed on a ResourceQ column. The columns were regenerated and equilibrated with an ABA wash, consisting of 3 CV of Buffer A1/2 followed by 3 CV of Buffer B1/2 and subsequently 3 CV of Buffer A1/2. The columns were then equilibrated with 3 CV of 1% Buffer B1 (ResourceQ column) or 10% Buffer B2 (Heparin column). The NaCl concentration was lowered by diluting the sample in Buffer A1/2. The sample was loaded using a 50 ml SuperLoop (GE Healthcare) in 1% Buffer B1 (PRK1 HR1a) or 10% Buffer B2 (PRK3 HR1ab and PRK3 HR1abc) at a flow rate of 1 ml/min. The column was washed to baseline with 1% Buffer B1 or 10% Buffer B2 at a flow rate of 1 ml/min. The protein was eluted with a Buffer B1/2 gradient (starting at 1% Buffer B1 or 10% Buffer B2 and ending at 50% Buffer B1/2) across 3 CV at a flow rate of 1 ml/min. 1 ml fractions were collected and analysed by SDS-PAGE.

2.2.6.4 Glutathione sepharose affinity purification

The column was equilibrated with 3 CV Buffer S1/2/3. The sample was loaded using a 5 ml loop at a flow rate of 1 ml/min. The flow-through was collected and the column was washed to baseline with Buffer S1/2/3 at a flow rate of 1 ml/min. The flow-through and wash containing unbound protein were pooled. The column was washed (and regenerated) with Buffer G at a flow rate of 1 ml/min to elute GST.

2.2.6.5 Purification buffers

The relevant purification and biophysical analysis buffers for each protein are listed in Table 2.12.

Protein	Buffers
PRK1 HR1a	MTPBS (High salt), Buffer A1, Buffer B1, Buffer S1, CD buffer 1
PRK1 HR1b	MTPBS (High salt), Buffer S1, CD buffer 1
PRK1 HR1c	MTPBS (High salt), Buffer S1, CD buffer 1
PRK1 HR1ab	MTPBS (High salt), Buffer S2, CD buffer 2
PRK1 HR1abc	MTPBS (High salt), Buffer S2, CD buffer 2
PRK3 HR1a	MTPBS (High salt), Buffer S1, CD buffer 1
PRK3 HR1b	MTPBS (High salt), Buffer S3, CD buffer 3
PRK3 HR1c	MTPBS (High salt), Buffer S1, CD buffer 1
PRK3 HR1ab	MTPBS (High salt), Buffer A2, Buffer B2, Buffer S3, CD buffer 3
PRK3 HR1abc	MTPBS (High salt), Buffer A2, Buffer B2, Buffer S3, CD buffer 3
RhoA Δ4 F25N Q63L, RhoA Δ12 F25N Q63L, Cdc42 Δ7 Q61L	TNMD, $TNMD_{250}$, Buffer S4

Table 2.12 Purification and biophysical analysis buffers for each protein.

2.2.6.6 Protein concentrating

For all intermediate steps of purifications, proteins were purified in an Amicon stirred concentrator cell (Millipore) with a YM-10 membrane (or YM-3 for single HR1 domains). The final concentrating was done in an Amicon Ultra-4 Centrifugal Filter Unit (Millipore) with an Ultracel-10 or Ultracel-3 membrane.

2.2.7 Determination of protein concentration

2.2.7.1 Absorbance

The theoretical extinction coefficient of proteins at 280 nm was predicted from the number of tryptophan and tyrosine residues in the sequence using Pepstats. For GTPases, the extinction coefficient of the guanine nucleotide, 7720 $M^{-1}cm^{-1}$, was also taken into account.

Three µl of protein were taken up in a glass capillary (Shimadzu) of 0.05 cm pathlength and the absorbance at 280 nm was measured within a linear range of 0.1-1.0 in a UV-1800 UV-VIS spectrophotometer (Shimadzu). The concentration was determined according to the Beer-Lambert Law,

$$A = \varepsilon c l$$

where A is the absorbance at 280 nm, ε is the theoretical extinction coefficient in $M^{-1}cm^{-1}$, c is the concentration in M and l is the pathlength of the capillary in cm.

2.2.7.2 Bradford assay

For proteins with few or no aromatic residues or proteins unstable at room temperature, the concentration was estimated using a Bradford reagent-based assay. 800 µl of SMQ were added to 200 µl BioRad reagent in a 1 ml cuvette. 2 µl of protein were then added and the contents of the cuvette were mixed by inversion. The absorbance reading at 595 nm was recorded after 3 min using a spectrophotometer. The absorbance reading of a blank sample (800 µl of SMQ and 200 µl BioRad reagent) was then subtracted from the absorbance of the protein sample. The recordings were taken in duplicates (within 5% of each other) and averaged. The subtracted, averaged reading was used to estimate the amount of protein, in µg, present in the cuvette based on an equation derived from the standard curve of a Bradford assay with IgG. Samples that gave spectrophotometer values larger than 1.0 (outside the linear range) were diluted and the assay was then repeated.

2.2.7.3 Amino Acid Analysis

Amino acid analysis was performed by the Protein and Nucleic Acid Chemistry facility at the Department of Biochemistry, University of Cambridge.

2.2.8 Mass spectrometry

MALDI-TOF was performed by the Protein and Nucleic Acid Chemistry facility at the Department of Biochemistry, University of Cambridge, and by the Proteomics facility at the Cambridge Centre for Proteomics, University of Cambridge.

2.2.9 Nucleotide exchange

2.2.9.1 GMPPNP exchange

The alkaline phosphatase agarose bead slurry was equilibrated with 10 mM Tris-HCl pH 7.5, 0.1 mM $ZnCl_2$.

For the nucleotide exchange, 500 µl GTPase (up to 1 mM), 333 µl 20 mM GMPPNP (Sigma-Aldrich), 333 µl 50% alkaline phosphatase agarose bead slurry and 91 µl 3 M $(NH_4)_2SO_4$ in 10 mM Tris-HCl pH 7.5 were mixed to a total volume of 1257 µl. The mixture was then split into two separate exchange reactions, which were left rotating at 37 ^{o}C for 4 h 30 min. After the exchange, 2x~600 µl (the supernatant) were spun down at 16060 x g for 5 min at 4 ^{o}C and then purified on a Superdex75 16/60 column, as described in section 2.2.6.2 to buffer exchange the protein and remove any aggregated protein and the excess GMPPNP.

2.2.9.2 HPLC analysis

50 µg of protein were mixed with 2 µl of 0.9 M perchlorate, shaken and made up to 100 µl with 0.6 M $NH_4H_2PO_4$ pH 4.0. The precipitated protein was pelleted at 16060 x g for 5 min and 100 µl of the supernatant were injected on a PartiSphere SAX column (125 mm x 4.6

mm) (Whatman) on a GP50 Gradient Pump HPLC system (Dionex). The column was run isocratically with 0.6 M $NH_4H_2PO_4$ pH 4.0 for 15 min at 1.5 ml/min. One hundred µl of 0.1 mM GDP, GTP and GMPPNP standards (Sigma-Aldrich) were individually run and compared to the guanosine nucleotide released by the exchanged protein.

2.2.9.3 [^{3}H-GTP] exchange

[8,5'-^{3}H]GTP (0.15 miCi) (GE Healthcare) in 75% EtOH was dried by centrifugal evaporation. At least 0.7 mg of GTPase were added to the dry radionucleotide and the volume was made up to 145 µl in SPA exchange buffer. 16 µl of 3 M $(NH_4)_2SO_4$ were then added and the reaction was incubated at 37 ^oC for 3 h. The reaction was quenched by adding 3 µl of 1 M $MgCl_2$. Excess radionucleotide was then removed using 2.5 ml G25 Sephadex spin columns (GE Healthcare) previously equilibrated in Buffer S4.

The concentration was determined using a Bradford assay and corrected using a correction factor obtained from amino acid analysis.

2.2.10 Circular dichroism

2.2.10.1 Sample preparation

Proteins were buffer exchanged into their respective CD buffer using a PD10 column and then were concentrated to 1 mg/ml. The proteins were further diluted with CD buffer to a concentration of 0.2 mg/ml prior to recording the experiments.

2.2.10.2 Data recording

The proteins were loaded in a 0.1 mm pathlength quartz cuvette which was placed in an AVIV 410 instrument (Aviv Biomedical). The ellipticity in millidegrees was measured at 18 ^oC between 185 nm and 260 nm at 1 nm intervals with a 1 s averaging time/point. Four scans were recorded in total and averaged. The same was repeated for each of the different CD

buffers. The buffer control data was then subtracted from the experimental data, which was followed by smoothing the plots.

The recorded data in millidegrees were converted to mean residue ellipticity by

$$MRE = \frac{\theta \times 0.1 \times MRW}{C \times l}$$

where MRE is the calculated mean residue ellipticity in degrees cm^2 $dmol^{-1}$ $residue^{-1}$, θ is the recorded ellipticity in millidegrees, l is the pathlength in cm, C is the protein concentration in mg/ml and MRW is the mean residue weight of each protein, calculated by

$$MRW = \frac{M_r}{n-1}$$

where M_r is the molecular mass of the protein in Da and n is the number of residues.

For thermal melts, ellipticity at 222 nm was monitored at an increasing temperature between 18 ^oC and 80 ^oC at 2 ^oC increments with a 30 sec equilibration time. Subsequently, refolding was monitored in a similar way by cooling the sample back to 18 ^oC.

2.2.10.3 Data analysis

The data were analysed using the DICHROWEB server and percentage helicity was estimated using the SELCON3 programme and reference set 3 (Sreerama & Woody, 1993; Whitmore & Wallace, 2008).

2.2.11 Analytical ultracentrifugation

The protein and reference sectors of standard double-sector Epon centrepieces with sapphire windows were filled with 400 µl of protein and 400 µl of matching buffer, respectively. The AUC cell centrepieces were loaded into an An60Ti four-hole rotor and placed in an Optima XL-I centrifuge (Beckman Coulter). For temperature sensitive protein samples, the AUC cells were first cooled down to 4 ^oC prior to sample loading. Experiments

were carried out at 4 ^{o}C and at a rotor speed of 40000 rpm. Interference optics data were recorded in the continuous mode without averaging and with radial increments of 0.003 cm. Scans were recorded at 1 min 30 sec intervals until no further sedimentation was observed. The systematic noise was then subtracted.

Estimates for the solvent density, solvent viscosity and the protein partial specific volume were calculated using the program SEDNTERP (Laue et *al.*, 1992). A total of 200 scans were analysed and the data were fit to a continuous c(s) distribution (with a Marquardt-Levenberg algorithm) using SEDFIT v14.1 (Schuck, 2000). A range of s-values from 0-10 was chosen first when fitting the data to confirm absence of very large species. Subsequently, a shorter range, e.g. 0-5 was chosen to allow for greater resolution in the c(s) distribution.

For the PRK1 HR1a/RhoA Δ4 F25N Q63L experiment, the proteins were mixed and incubated on ice for 30 min before being loaded into the AUC cells.

Appendix 3 lists all the AUC experiments performed, the parameters estimated by SEDNTERP, the number of scans recorded and the parameters determined.

2.2.12 NMR Spectroscopy

2.2.12.1 Expression of isotopically-labelled proteins

For small scale expression trials, 10 ml 2TY with the appropriate antibiotic(s) were inoculated and grown overnight at 37 ^{o}C shaking at 180 rpm. The overnight culture was then resuspended in either 10 ml MOPS or M9 minimal medium and 1 ml was then added to 10 ml MOPS or M9 minimal medium (x3 for Uninduced, Induced at 37 ^{o}C and Induced at 20 ^{o}C).

For large scale protein expression, 20 ml 2TY cultures with the appropriate antibiotics were grown overnight at 37 ^{o}C. The cultures were pelleted and resuspended in 10 ml MOPS or M9 and then added to 500 ml of MOPS or M9 minimal medium with the appropriate antibiotic. ^{15}N-NH_4Cl was also added to a final concentration of 1 mg/ml as well as 0.4% unlabelled glucose (for ^{15}N-labelled proteins) or 0.3% labelled glucose (for ^{15}N/^{13}C-labelled proteins). The cultures were then induced, grown and processed as for unlabelled proteins.

2.2.12.2 NMR experiments

NMR spectra were recorded as follows.

Protein	Conc. (mM)	Buffer	Experiment	Temperature (K)	Titration points	Spectrometer
PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	2D NOESY Mixing time: 150 msec	298	-	Bruker DRX500
PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	2D TOCSY Mixing time: 60 msec	298	-	Bruker DRX500
^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC (full)	298	-	Bruker DRX500
^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC (folded)	298	-	Bruker DRX500
^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-separated NOESY Mixing time: 150 msec	298	-	Bruker DRX500
^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-separated TOCSY Mixing time: 60 msec	298	-	Bruker DRX500
^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC (t1)	298	Delay times (msec): 10, 50, 100, 150, 250, 400, 600, 800	Bruker DRX500
^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC (t2)	298	Delay times (msec): 14.4, 28.8, 43.2, 57.6, 72, 86.4, 100.8	Bruker DRX500

^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC (hNOE reference)	298	-	Bruker DRX500
^{15}N PRK1 HR1c	1.6	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC (hNOE saturated)	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	HNCA	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	HNCACB	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	HN(CO)CA	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	HN(CO)CACB	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{13}C-HSQC (full)	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{13}C-HSQC (folded)	298	-	Bruker DRX500

^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	HCCH-TOCSY Mixing time: 18 msec	298	-	Bruker DRX500
^{15}N, ^{13}C PRK1 HR1c	1.2	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{13}C-separated-NOESY Mixing time: 100 msec	298	-	Avance AV800
PRK3 HR1c	0.7	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	2D NOESY Mixing time: 150 msec	298	-	Bruker DRX500
PRK3 HR1c	0.7	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	2D TOCSY Mixing time: 60 msec	298	-	Bruker DRX500
PRK1 HR1abc	0.45	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 5 mM DTT, 0.05% NaN_3, 10% D_2O	2D NOESY Mixing time: 150 msec	278	-	Bruker DRX500
PRK1 HR1abc	0.45	20 mM sodium phosphate pH 7.3, 150 mM NaCl, 5 mM DTT, 0.05% NaN_3, 10% D_2O	2D NOESY Mixing time: 150 msec	288	-	Bruker DRX500
^{15}N PRK1 HR1c + HR1a	0.3	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	1:0 1:0.25 1:0.5 1:1 1:2 1:6	Bruker DRX500
^{15}N PRK1 HR1c + HR1a + HR1b	0.3	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	1:0.5 1:1 1:2 1:6	Avance AV800

^{15}N PRK1 HR1c + HR1b	0.3	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	1:0 1:0.25 1:0.5 1:1 1:6	Avance AV800
^{15}N PRK1 HR1c + HR1b + HR1a	0.3	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	1:6:0 1:6:0.5 1:6:1 1:6:2 1:6:6	Avance AV800
^{15}N PRK1 HR1c + HR1ab	0.3	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	1:0 1:0.25 1:0.5 1:1 1:2 1:6	Avance AV800
^{15}N PRK1 HR1ab	0.357 0.2 0.1 0.05 0.015	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	-	Avance AV800
^{15}N PRK1 HR1ab	0.05	50 mM Tris-HCl pH 7.5, 250 mM NaCl, 5 mM DTT, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	-	Avance AV800
^{15}N PRK1 HR1ab + HR1c	0.05	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT, 0.05% NaN_3, 10% D_2O	^{15}N-HSQC	298	1:0 1:0.5 1:1 1:2 1:6	Avance AV800

Table 2.13 NMR experiments.

2.2.12.3 NMR data processing

NMR data were processed using Azara (Wayne Boucher, University of Cambridge) with the help of Dr. Helen Mott.

2.2.12.4 NMR spectra analysis

NMR spectra were analysed using CCPN Analysis v2.3 (Vranken *et al.*, 2005).

2.2.12.5 Structure calculations

Structure calculations were performed using ARIA 2.3.1 (Rieping *et al.*, 2007) interfaced to CNS 1.3 (Brünger *et al.*, 1998). The predicted dihedral angles from TALOS and the intensities of peaks in ^{15}N-separated NOESY/^{13}C-separated NOESY spectra were used to derive dihedral angle and distance restraints, respectively. Chemical shift tolerances were set at 0.04 ppm for the indirect dimension (proton 1), 0.02 ppm for the direct dimension (proton 2) and 0.5 ppm for the heteronuclear dimensions. Structures were calculated iteratively for a total of 8 iterations. During iteration 8, 100 structures were calculated and of these the 50 with the lowest energy were refined in water. The final ensemble of structures consisted of the 35 lowest energy structures.

2.2.12.6 Chemical shift mapping

Chemical shift changes were calculated using the following equation

$$\delta = \sqrt{\delta_{1_H}^2 + (0.15\delta_{15_N})^2}$$

where δ is the combined shift change and δ_{1_H} and δ_{15_N} are the chemical shift changes in the ^{1}H and ^{15}N dimensions, respectively. Residues which could not be reliably assigned due to overlap were excluded from the analysis.

2.2.13 Crystallisation trials

Purified PRK1 HR1abc protein was concentrated to 10 mg/ml. One to one (100 nl + 100 nl) and two to one (200 nl + 100 nl) protein to buffer drops were set up in seven different 96-well sitting drop crystallisation plates: Classics Lite Suite (QIAGEN), JCSG+ (QIAGEN), PACT (Molecular Dimensions), Wizard I&II (Molecular Dimensions), Wizard III&IV (Molecular Dimensions), PEGI (QIAGEN) and PEG II (QIAGEN) using the MOSQUITO CRYSTAL robot (TTPLabtech).

Custom crystallisation plates were set up using the DRAGONFLY (TTPLabtech) and are listed in the table below. Crystallisation trials were set up as before.

Plate 1	PEG3350 (10-35 mM), NaF (10-300 µM)
Plate 2	PEG3350 (10-35 mM), KI (10-300 µM)
Plate 3	PEG4000 (10-35 mM), $(NH_4)_2SO_4$ (10-300 µM)
Plate 4	PEG4000 (10-30 mM), $MgSO_4$ (10-250 µM)
Plate 5	PEG3350 (10-35 mM), Na_2SO_4 (10-300 µM)

Table 2.14 Custom crystallisation plates.

For the RhoA Δ12 F25N Q63L/PRK1 HR1abc co-crystallisation trials, the proteins were mixed at a 1:1 ratio with a final concentration of 250 mM for each protein and were incubated on ice for 30 min before setting up the commercial crystallisation plates as with PRK1 HR1abc.

2.2.14 Isothermal Titration Calorimetry

Proteins were all buffer exchanged using PD10 columns into the same batch of ITC buffer. ITC experiments were carried out on an iTC_{200} microcalorimeter (MicroCal) (GE Healthcare). The cell was filled with 200 µl of buffer or protein and the syringe was filled with 100 µl of protein. The concentration of protein placed in the syringe was ten-fold that of the concentration of protein in the cell. Protein into protein, as well as buffer into buffer and protein into buffer control experiments were recorded. Each experiment was performed at 25 ^{o}C and comprised a total of 19 injections over 50 min; the first injection was of 0.4 µl and the remaining 18 injections of 2 µl. Protein into buffer control data were subtracted from protein into protein data, which were then fit with Origin using a one-site binding model to determine the K_d.

2.2.15 Cross-linking

2.2.15.1 Cross-linking with EDC

Proteins were buffer exchanged into EDC buffer using a PD10 column. Cross-linking reactions were set up at room temperature or at 4 ^{o}C according to the table. 10 µl samples were taken, 10 µl 2x SDS Loading Buffer was added and 10 µl were analysed by SDS-PAGE.

SDS-PAGE trials				
Protein	**[Protein] (mM)**	**[EDC] (mM)**	**Volume (µl)**	**Intervals at which samples were taken**
PRK1 HR1a	1.5	10, 20, 40	50	30 min, 5 h, overnight (4 ^{o}C)
PRK1 HR1ab	1.5	10, 20, 40	50	30 min, 5 h, overnight (4 ^{o}C)
PRK1 HR1abc	0.7	10, 20, 40	50	30 min, 5 h, overnight (4 ^{o}C)
PRK1 HR1ab + PRK1 HR1c	1.5 (HR1ab) + 1.5 (HR1c)	10	50	Overnight (4 ^{o}C)

Table 2.15 Cross-linking reactions with EDC.

2.2.15.2 Cross-linking with DMS/DMP

Proteins were buffer exchanged into DM buffer using a PD10 column. Cross-linking reactions were set up at room temperature or at 4 ^{o}C according to the table. 10 µl samples were taken, 10 µl 2x SDS Loading Buffer was added and 10 µl were analysed by SDS-PAGE.

SDS-PAGE trials				
Protein	**[Protein] (mM)**	**[DMS]/[DMP] (mM)**	**Volume (µl)**	**Intervals at which samples were taken**
PRK1 HR1a	1.5	10, 15, 20	50	30 min, 5 h, overnight (4 ^{o}C)
PRK1 HR1ab	1.5	10, 15, 20	50	30 min, 5 h, overnight (4 ^{o}C)
PRK1 HR1abc	0.7	10, 15, 20	50	30 min, 5 h, overnight (4 ^{o}C)
PRK1 HR1ab + PRK1 HR1c	1.5 (HR1ab) + 1.5 (HR1c)	20	50	Overnight (4 ^{o}C)

Table 2.16 Cross-linking reactions with DMS/DMP.

2.2.16 Pull-downs

Two litres of *E.coli* culture expressing the appropriate protein were spun down, lysed and purified as detailed in section 2.2.5. The supernatant was added to 4 ml of 50% glutathione agarose bead slurry to saturate the beads with the GST-tagged GTPase. Ten µl of the purified, glutathione agarose-immobilised GST-tagged GTPase were also taken to assess nucleotide state by HPLC as described in section 2.2.9.2.

The beads were then washed in TNMD buffer to reduce the NaCl concentration to 150 mM. 200 µl of 50% bead slurry was aliquoted in 1.5 ml Eppendorf tubes and HR1 effector proteins were added to a final concentration of 200 µM. The tubes were then rotated on a wheel at 4 ^{o}C for 30 min to facilitate mixing and binding. The reactions were then spun at 1000 x g for 10 min, the supernatant was removed and a gel sample was taken. TNMD buffer to equal volume (200 µl) was added to the beads to wash off unbound effector protein. The beads were spun at 1000 x g for 10 min and a gel sample of the wash was taken. The beads were washed for a total of seven times and were finally resuspended in 50% slurry (200 µl) with TNMD buffer. A gel sample of the beads (pull-down sample) was taken. Ten µl of each sample (a 1 in 2 dilution in 2x SDS Loading Buffer) were then analysed by SDS-PAGE.

2.2.17 Scintillation Proximity Assay

Scintillation proximity assays (SPAs) were set up as competition assays. 30 nM GST-ACK or 200 nM PRK1 HR1a-His_6 and equimolar Cdc42 Δ7 Q61L·[^{3}H]GTP or RhoA Δ7 F25N Q63L·[^{3}H]GTP, respectively, were added to SPA reaction mix so that GTPase/effector complexes were bound to SPA fluoromicrospheres via anti-GST or anti-His antibodies. Twelve-fold serial dilutions of competitor protein (PRK3 HR1c or PRK1 HR1abc) were set up and 25 µl were added to 175 µl of SPA reaction mix to give the final competitor concentrations shown in each experiment. Control experiments with untagged ACK and

GST-PRK1 HR1a as competitors were also performed. GST-ACK, PRK1 HR1a-His_6, untagged ACK and GST-PRK1 HR1a were a kind gift from Dr. Darerca Owen.

The plates were incubated on a shaking platform for 45 min at 18 ^{o}C and then spun at 1000 x g for 2 min to pellet the beads. Each well was counted for 5 minutes in a MicroBeta Trilux scintillation counter (PerkinElmer). Non-linear regression analysis was performed using GraFit5 according to the following equation:

$$\text{SPA Signal} = \frac{[S_{max}[(K_i\text{P} + K_d I_0) - \sqrt{(K_i\text{P} + S_0 + K_d I_0)^2 - 4K_i^2 E_0 S_0}]]}{K_i(P - \sqrt{P^2 - 4E_0 S_0})}$$

where S_{max} is the maximum SPA signal, K_d is the dissociation constant between the tagged effector and the GTPase, S_0 is the total concentration of labelled GTPase, K_i is the equilibrium dissociation constant between the GTPase and the untagged effector, E_0 is the total concentration of tagged effector protein and P is (E_0 + S_0 + K_d).

2.2.18 Computational methods

The software and software version used are listed in Table 2.17.

Software	Task
PyMOL v1.3	Structure analysis and structure images
Clustal Omega	Sequence alignment
Pepstats	Protein pI estimation
ÄKTA Purifier with Unicorn 5 (GE Healthcare)	Protein chromatography
Jpred (Drozdetskiy *et al.*, 2015)	Secondary structure prediction
Dichroweb server	CD data analysis
SEDNTERP (Laue *et al.*, 1992)	AUC parameter estimation
SEDFIT v14.1 (Schuck, 2000)	AUC data analysis
GUSSI	AUC profile overlay
HYDROPRO	Frictional ratio prediction

CCPN Analysis v2.4 (Vranken *et al.*, 2005)	NMR spectra assignment
Azara (Wayne Boucher, University of Cambridge)	NMR data processing
ChemDraw v17.1	Drawing figures
ARIA v2.3.1	Structure calculations
Origin	ITC data analysis
NACCESS (Hubbard & Thornton, 1993)	Determination of solvent accessibility
HADDOCK	Protein docking
Dionex with Chromeleon v6.50 (Thermo Fischer)	HPLC analysis
Chromas Lite v2.1.1	Analysis of DNA sequences
Microsoft Office 10	Drawing figures
GraFit5	SPA data analysis

Table 2.17 Software.

Chapter 3 Purification of PRK1 and PRK3 HR1 constructs and investigation of their oligomeric state

3.1 Introduction

The starting point of this work was to design and clone PRK1-PRK3 HR1 constructs. The proteins would be expressed as fusions with an N-terminal GST tag to aid solubility and purification. Optimal conditions for expression and solubility were identified in small scale expression trials and then the purification of each protein was separately optimised to yield protein with high purity for biophysical and structural analysis.

3.2 Protein purification

3.2.1 Construct cloning

Existing but also new constructs were designed by predicting secondary structures in jpred (Drozdetskiy *et al.*, 2015) and choosing appropriate construct limits. To express N-terminal GST fusions of PRK1 HR1a and PRK1 HR1ab, pGEX6P constructs were used which were already available in the lab.

PRK1 HR1c, PRK2 HR1c, PRK2 HR1bc, PRK2 HR1abc, PRK3 HR1c, PRK3 HR1bc and PRK3 HR1abc were cloned into pGEXHISP during a Master's project in this lab (Sophocleous, 2014), while pGEXHISP constructs of PRK1 HR1b, PRK3 HR1a, PRK3 HR1b and PRK3 HR1ab were already available in the lab. To express GST fusions of these HR1 constructs, a stop codon was introduced by mutagenesis of pGEXHISP constructs (to remove the C-terminal His_6 tag).

PRK1 HR1bc and PRK1 HR1abc constructs were obtained by sub-cloning into pCR2.1(TOPO) vector, followed by *Bam*HI/*Eco*RI ligation into pGEXHISP and then introducing a stop codon in the pGEXHISP vector in the same way as the other HR1 constructs.

All the GST fusions carry an HRV 3C protease cleavage site between GST and the HR1 construct to allow removal of the tag after the initial affinity purification step.

3.2.2 Expression trials

To assess the expression and solubility of the cloned fusion proteins, expression trials were carried out in BL21 and Rosetta2 strains. Expression at two temperatures was tested; 5 h at 37 ^{o}C or 20 h at 20 ^{o}C (both vectors).

In the expression trials, all but one PRK1 domain tested showed increased solubility at 20 ^{o}C than at 37 ^{o}C (most likely because they can more slowly fold and thus do not go into inclusion bodies), and expression was increased in the Rosetta2 strain (Figures 3.1A-D; 3.1F). The increased expression in Rosetta2 cells should be of no surprise as PRK1 contains 32 rare codons; 14 in HR1a, 1 in the linker between HR1a and HR1b, 3 in HR1b and 14 in HR1c. GST-PRK1 HR1bc was totally insoluble at both temperatures in both strains (Figure 3.1E).

The PRK2 HR1 domains tested showed poor solubility at both 37 ^{o}C and 20 ^{o}C (Figures 3.2A-C) and significant breakdown. At this stage it was decided that PRK2 domains would not be further studied.

Four out of six PRK3 HR1 domains were soluble at 37 ^{o}C and showed improved solubility when expressed at 20 ^{o}C (Figures 3.3A-C; 3.3F). GST-PRK3 HR1ab was only soluble at 20 ^{o}C (Figure 3.3D). GST-PRK3 HR1bc was soluble at 20 ^{o}C (with some breakdown) but this di-domain construct was not further studied as a result of the lack of solubility of its PRK1 homolog (Figure 3.3E). GST breakdown could be observed for most PRK3 constructs. There was also a slight increase in expression in Rosetta2 for some PRK3 HR1 domains (Figures 3.3A-B; 3.3D).

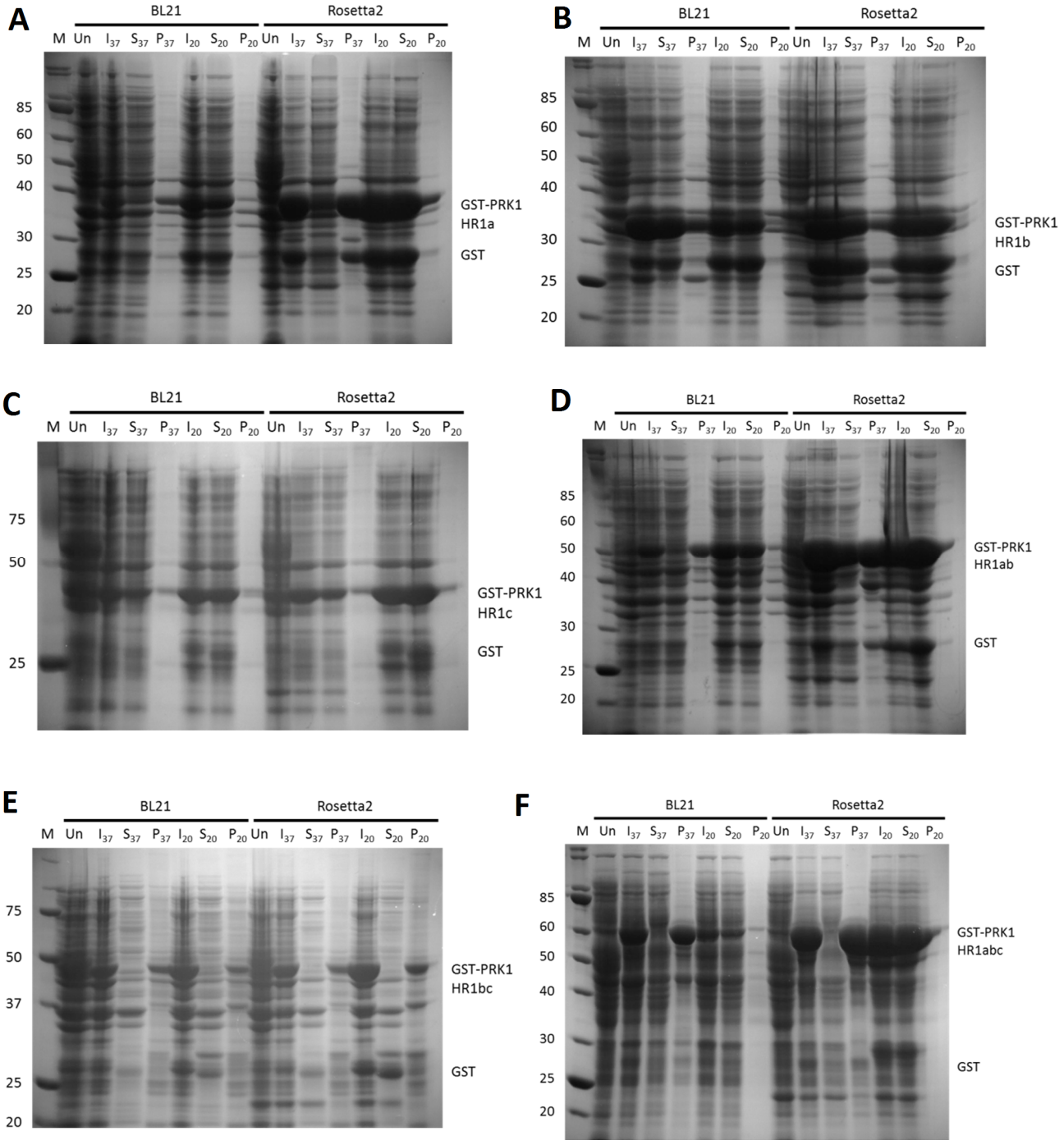


Figure 3.1 SDS-PAGE Analysis of PRK1 construct expression trials. Solubility of fusion protein was assessed by identifying band of approximate molecular weight in soluble or pellet fractions. Un for Uninduced sample, In for Induced (Total), S for Induced (Soluble) and P for Induced (Pellet). 20 and 37 denote the temperature at which the expression was carried out at. Laemmli gels were 12%. (**A**) GST-PRK1 HR1a (**B**) GST-PRK1 HR1b (**C**) GST-PRK1 HR1c (**D**) GST-PRK1 HR1ab (**E**) GST-PRK1 HR1bc (**F**) GST-PRK1 HR1abc

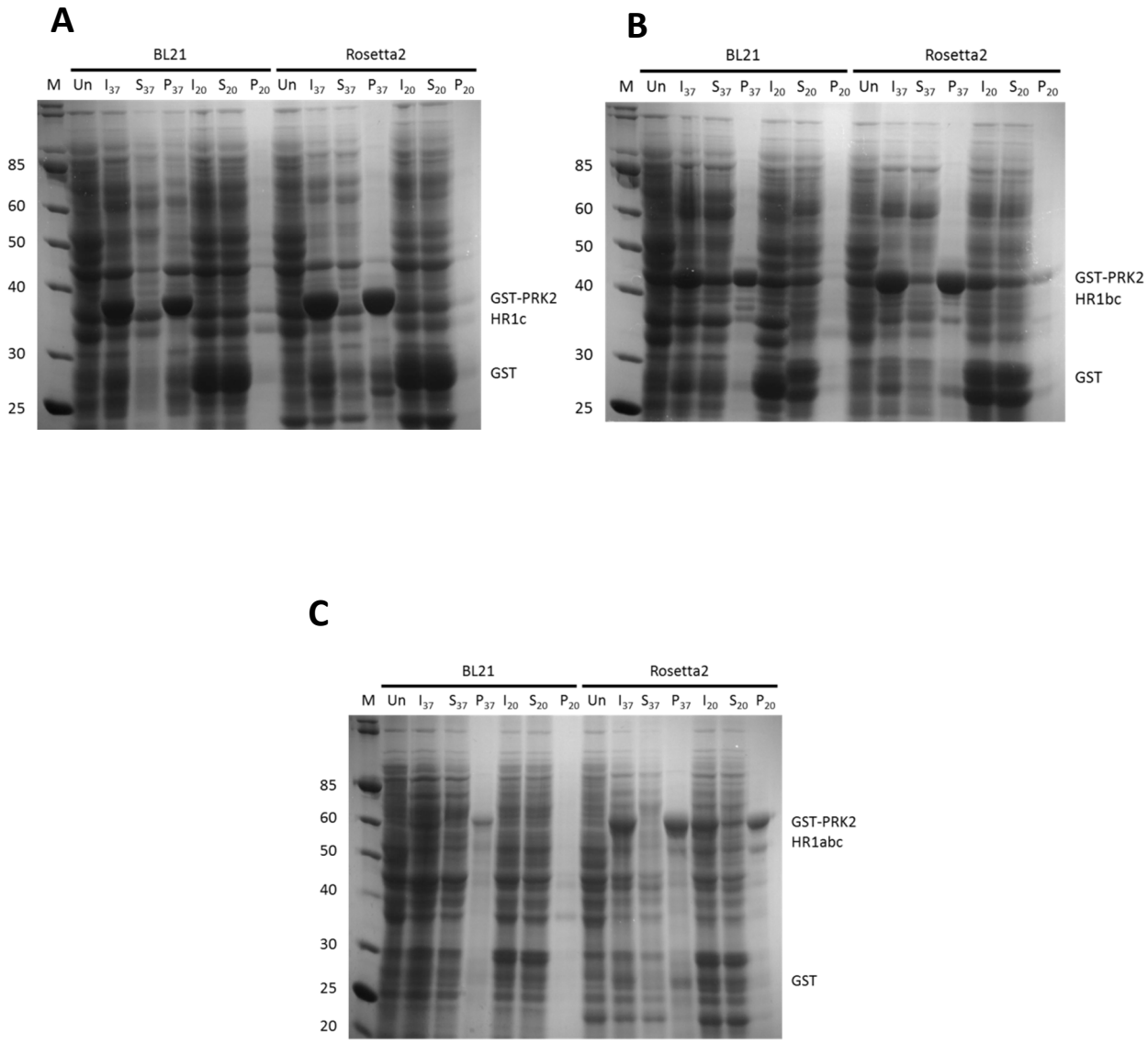


Figure 3.2 SDS-PAGE Analysis of PRK2 construct expression trials. Solubility of fusion protein was assessed by identifying band of approximate molecular weight in soluble or pellet fractions. Un for Uninduced sample, In for Induced (Total), S for Induced (Soluble) and P for Induced (Pellet). 20 and 37 denote the temperature at which the expression was carried out at. Laemmli gels were 12%. (**A**) GST-PRK2 HR1c (**B**) GST-PRK2 HR1bc (**C**) GST-PRK2 HR1abc

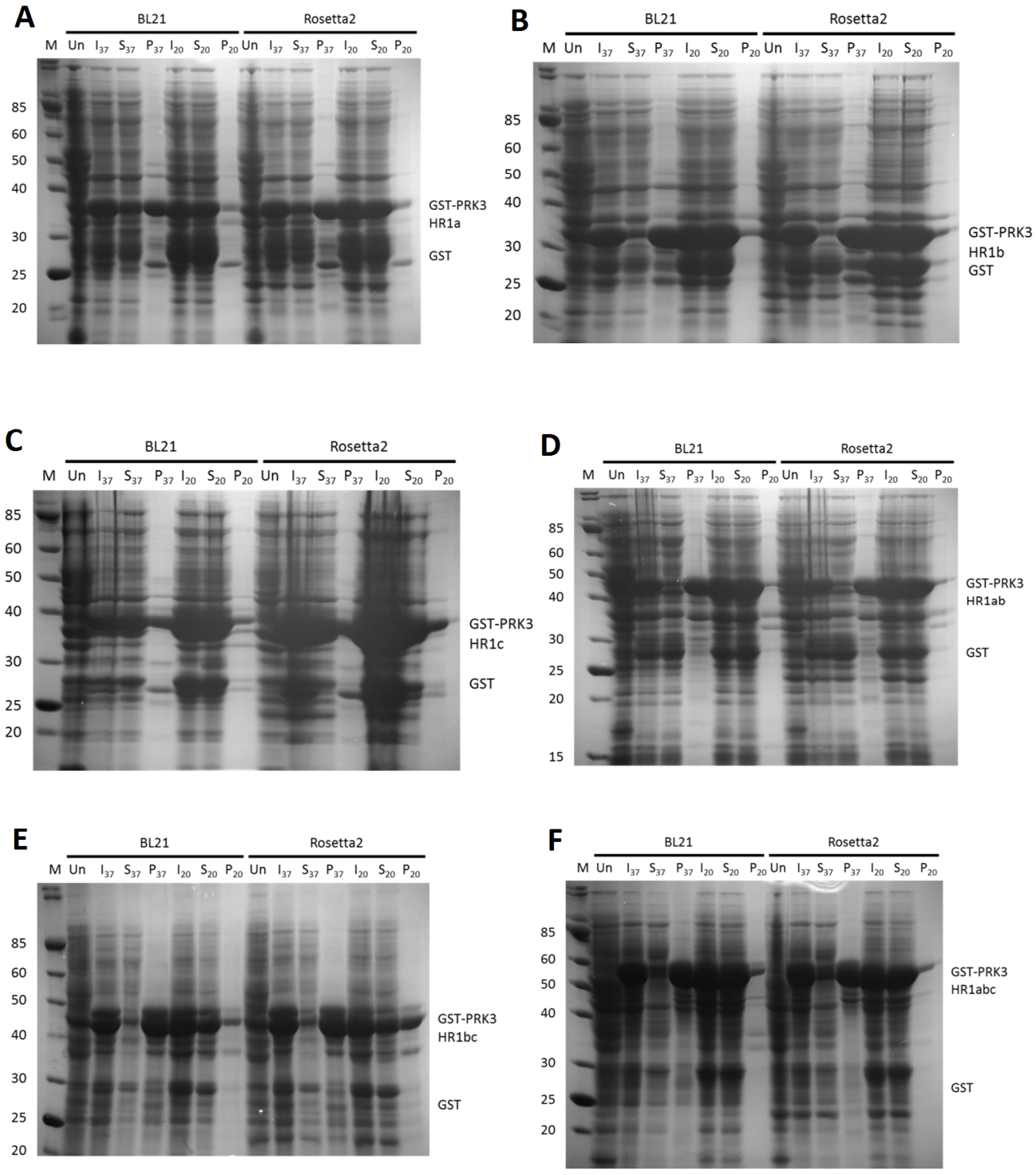


Figure 3.3 SDS-PAGE Analysis of PRK3 construct expression trials. Solubility of fusion protein was assessed by identifying band of approximate molecular weight in soluble or pellet fractions. Un for Uninduced sample, In for Induced (Total), S for Induced (Soluble) and P for Induced (Pellet). 20 and 37 denote the temperature at which the expression was carried out at. Laemmli gels were 12%. (**A**) GST-PRK3 HR1a (**B**) GST-PRK3 HR1b (**C**) GST-PRK3 HR1c (**D**) GST-PRK3 HR1ab (**E**) GST-PRK3 HR1bc (**F**) GST-PRK3 HR1abc

Table 3.1 summarises the expression trial results.

Construct	BL21		Rosetta2	
	37 °C	20 °C	37 °C	20 °C
PRK1 HR1a (1-106)	✓	✓✓	✓	✓✓✓
PRK1 HR1b (122-199)	✓	✓✓	✓✓	✓✓✓
PRK1 HR1c (201-297)	✓	✓✓	✓	✓✓✓
PRK1 HR1ab (1-199)	✓	✓✓	✓✓	✓✓✓
PRK1 HR1bc (122-297)	×	×	×	×
PRK1 HR1abc (1-297)	×	✓	×	✓✓✓
PRK2 HR1c (207-302)	×	×	×	×
PRK2 HR1bc (132-302)	✓	✓	✓	✓
PRK2 HR1abc (1-302)	×	×	×	✓
PRK3 HR1a (1-80)	✓	✓✓	✓	✓✓✓
PRK3 HR1b (100-171)	×	✓✓	×	✓✓✓
PRK3 HR1c (169-263)	✓	✓✓	✓	✓✓
PRK3 HR1ab (1-171)	×	✓✓	×	✓✓✓
PRK3 HR1bc (100-263)	✓	✓✓	✓	✓✓
PRK3 HR1abc (1-263)	×	✓✓	×	✓✓

Table 3.1 Summary of expression trials. The boxes shaded in yellow represent the best expression conditions carried forward in large scale expression. The number of ✓s represents the amount of soluble protein and a × represents complete lack of soluble protein.

3.2.3 Purification of GST HR1 fusions using glutathione agarose beads

The best expression conditions for each of the constructs identified from the expression trials were used for large scale expression and purification (4-6 L *E. coli* culture). GST fusions of all HR1 domain constructs were firstly purified using glutathione agarose beads. An example purification of GST-PRK1 HR1c is shown in Figure 3.4A.

3.2.4 Removal of GST tag by proteolytic cleavage

The N-terminal GST was removed by the addition of HRV 3C protease. Cleaved PRK1 HR1c eluted from the beads is shown in Figure 3.4B. 'High salt' MTPBS pH 7.3 was used during the glutathione agarose purification and post-cleavage washes to reduce non-specific binding of contaminants to the protein and protein precipitation. The supernatant and washes post-cleavage were then pooled and 5 mM EDTA was added to reduce any potential protein degradation. Furthermore, if the protein contained cysteines, 5 mM DTT was added.

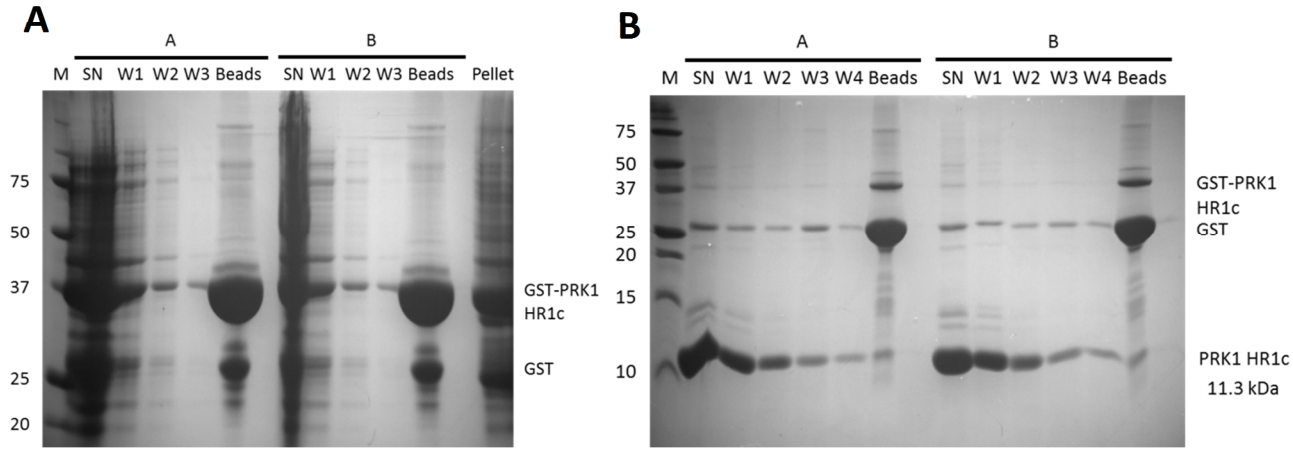


Figure 3.4 SDS-PAGE analysis of recombinant PRK1 HR1c protein expression and purification. Samples were taken at each step and analysed by SDS-PAGE. (**A**) GST-PRK1 HR1c immobilised on glutathione agarose beads. Gel was 12%. SN for supernatant, W1-3 for the 3 washes, Beads for 50% bead slurry and Pellet for insoluble fraction. (**B**) PRK1 HR1c washes post-cleavage. Gel was 18%. SN for supernatant and W1-4 for the 4 washes post-cleavage.

3.2.5 Further purification of PRK1 HR1b, PRK1 HR1c, PRK1 HR1abc and PRK3 HR1c by gel filtration

PRK1 HR1b was concentrated and loaded onto a Superdex 75 column. A single peak corresponding to PRK1 HR1b was observed (Figure 3.5A). Fractions were analysed by SDS-PAGE and C11-D2 were pooled and concentrated (Figure 3.5B) and the concentration was estimated using absorbance at 280 nm and a Bradford assay. HR1b contains only four arginines and no tryptophans, and as a result stains poorly with Coomassie dyes.

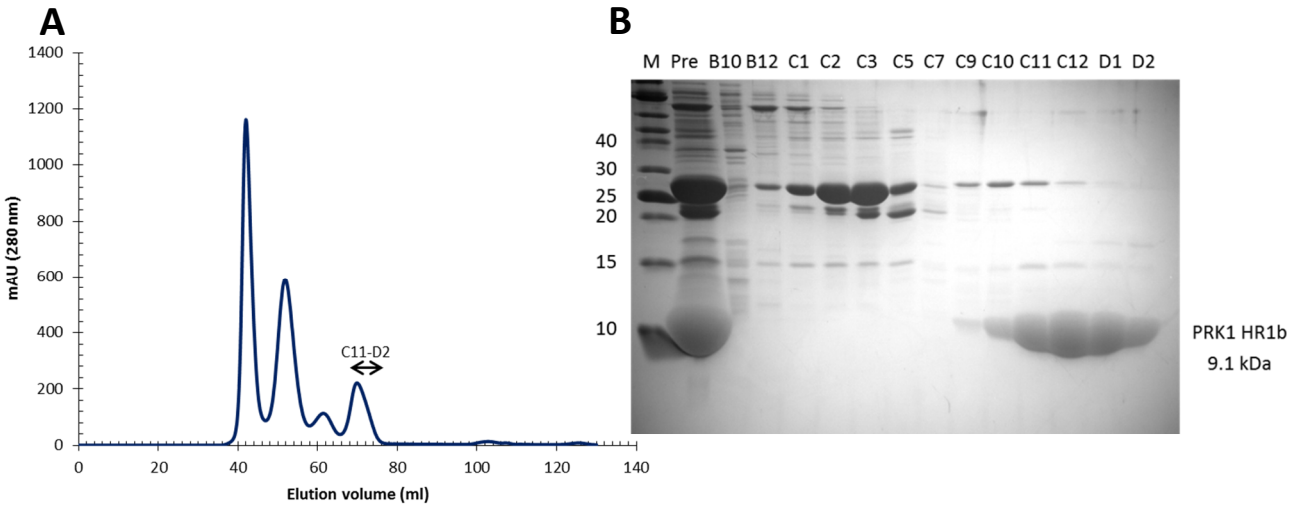


Figure 3.5 Further purification of PRK1 HR1b and corresponding SDS-PAGE analysis. (**A**) Gel filtration chromatogram of PRK1 HR1b on an S75 column. Blue line indicates A_{280}. (**B**) SDS-PAGE analysis of chromatogram in (**A**) on an 18% Laemmli gel. 'Pre' for sample loaded on the column. The double arrows indicate the fractions pooled.

PRK1 HR1c was also loaded onto a Superdex 75 column. HR1c lacks aromatic residues, so absorbance was alternatively monitored at a wavelength of 230 nm (peptide bond). A single peak corresponding to PRK1 HR1c was observed (Figure 3.6A). Fractions were analysed by SDS-PAGE and C8-C11 were pooled and concentrated (Figure 3.6B) and the concentration was estimated using a Bradford assay.

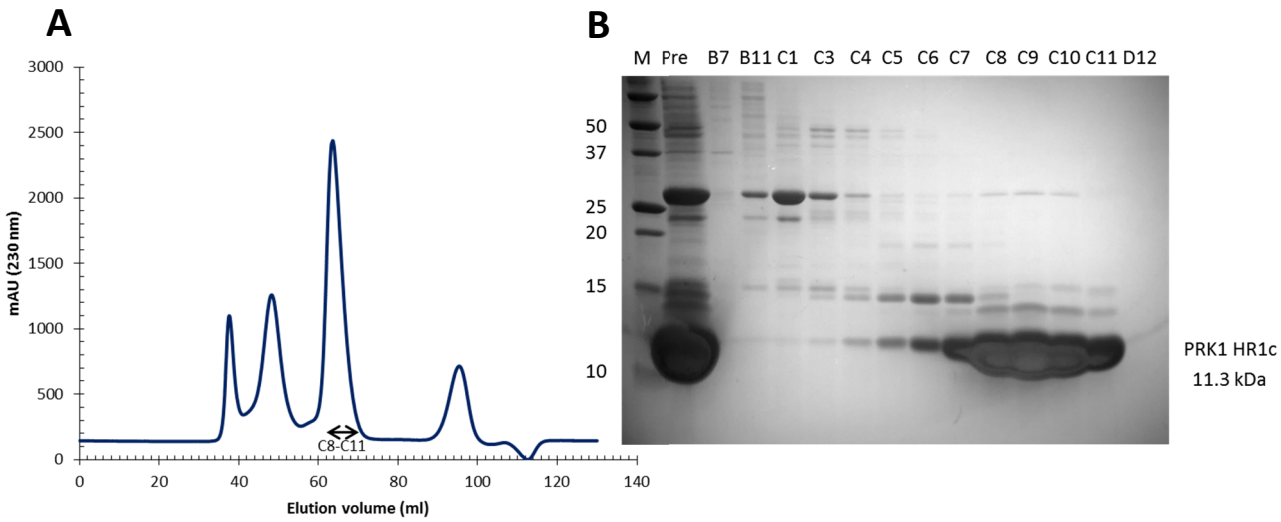


Figure 3.6 Further purification of PRK1 HR1c and corresponding SDS-PAGE analysis. (**A**) Gel filtration chromatogram of PRK1 HR1b on an S75 column. Blue line indicates A_{230}. (**B**) SDS-PAGE analysis of chromatogram in (**A**) on an 18% Laemmli gel. 'Pre' for sample loaded on the column. The double arrows indicate the fractions pooled.

PRK1 HR1abc was concentrated and the protein was loaded onto a Superdex 200 column. Absorbance was monitored at 280 nm. More than two peaks corresponding to HR1abc were observed (Figure 3.7A). Fractions were analysed by SDS-PAGE and C2-C12 were pooled (Figure 3.7B). Fractions D1-D8, which contained GST contamination, were pooled separately and loaded a second time on the Superdex 200 column (Figure 3.7C). Fractions C3-C12 (Figure 3.7D) were pooled together with fractions C1-C12 from the first run, concentrated, and the concentration was estimated using a Bradford assay.

PRK3 HR1c was also purified by two rounds of gel filtration. The protein was loaded onto a Superdex 75 column (Figure 3.8A). PRK3 HR1c also lacks aromatic residues, so absorbance was again monitored at a wavelength of 230 nm. One peak corresponding to HR1c was observed, with elution volume of 54 ml. This was in contrast to PRK1 HR1c which had an elution volume of 64 ml, indicating that PRK3 HR1c may be oligomerising in solution. Fractions were analysed by SDS-PAGE and C8-C10 were pooled (Figure 3.8B). Fractions C3-C7, which contained GST contamination, were pooled separately and loaded onto a second time on the Superdex 75 column (Figure 3.8C). Fractions C3-C10 (Figure 3.8D) were pooled together with fractions C8-C10 from the first run, concentrated, and the concentration was estimated using a Bradford assay.

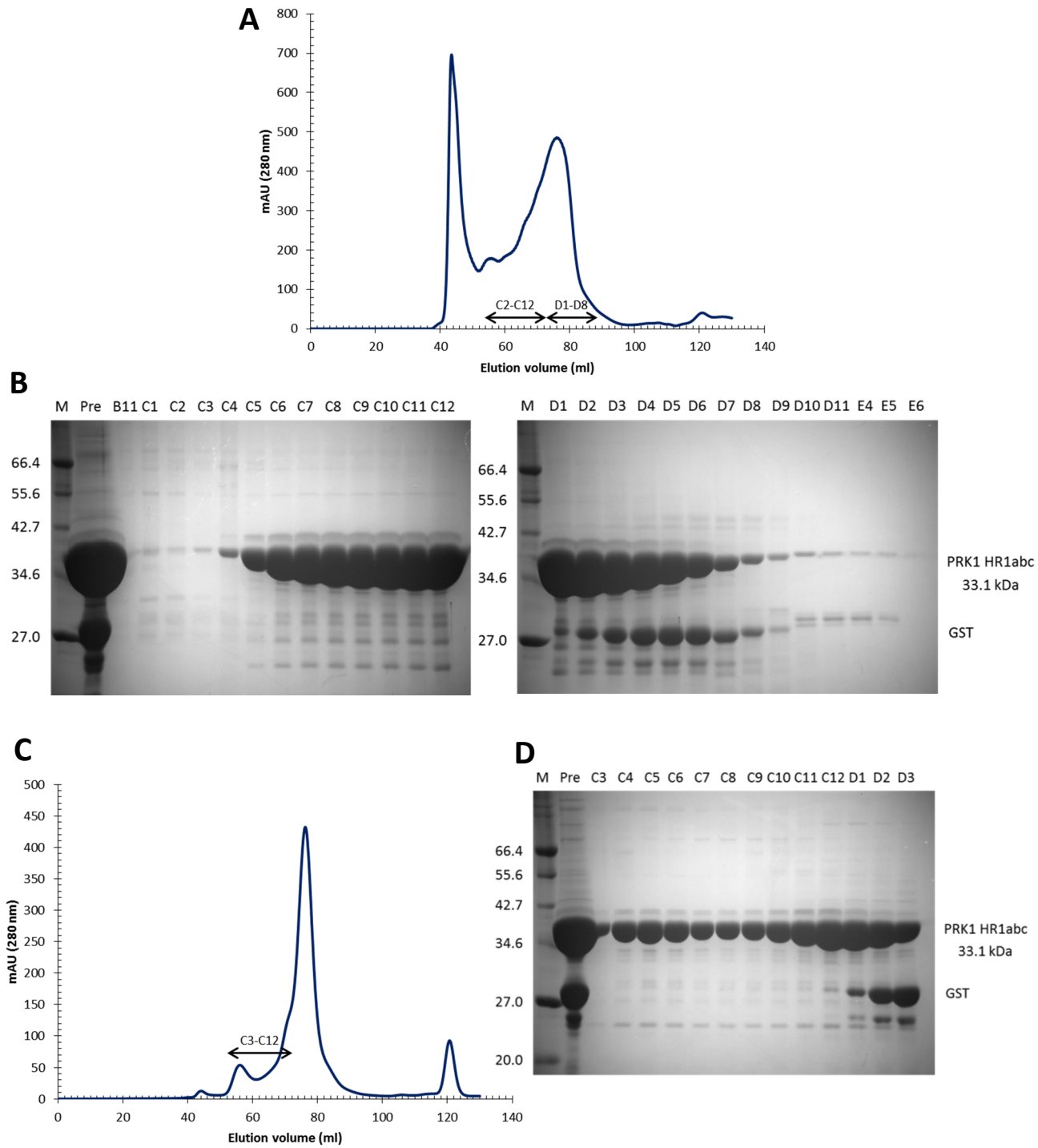


Figure 3.7 Further purification of PRK1 HR1abc and corresponding SDS-PAGE analysis. (**A**) Gel filtration chromatogram of PRK1 HR1abc on an S200 column. Blue line indicates A_{280}. (**B**) SDS-PAGE analysis of chromatogram in (**A**) on 12% Laemmli gels. (**C**) Second round gel filtration chromatogram of PRK1 HR1abc on an S200 column. Blue line indicates A_{280}. (**D**) SDS-PAGE analysis of chromatogram in (**C**) on a 12% Laemmli gel. 'Pre' for sample loaded on the column. The double arrows indicate the fractions pooled.

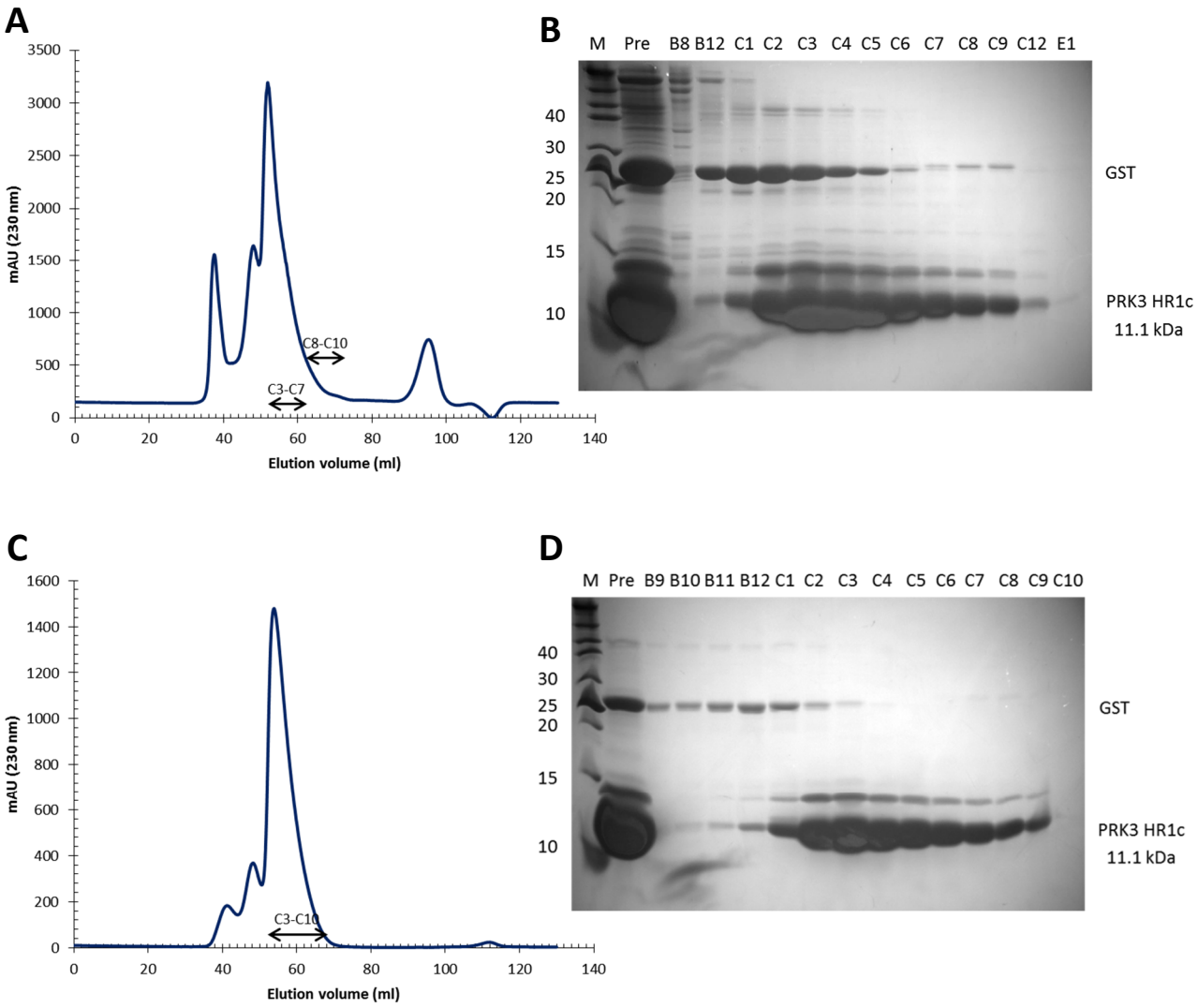


Figure 3.8 Further purification of PRK3 HR1c and corresponding SDS-PAGE analysis. (**A**) Gel filtration chromatogram of PRK3 HR1c on an S75 column. Blue line indicates A_{230}. (**B**) SDS-PAGE analysis of chromatogram in (**A**) on an 18% Laemmli gel. (**C**) Second round gel filtration chromatogram of PRK3 HR1c on S75 column. Blue line indicates A_{230}. (**D**) SDS-PAGE analysis of chromatogram in (**C**) on an 18% Laemmli gel. 'Pre' for sample loaded on the column. The double arrows indicate the fractions pooled.

3.2.6 Further purification of PRK1 HR1ab, PRK3 HR1a and PRK3 HR1b on a glutathione sepharose column followed by gel filtration

Since PRK1 HR1ab and GST have similar molecular weights and cannot be resolved by SDS-PAGE, it was difficult to tell whether there was any GST contamination. A single gel filtration purification step at this stage may not have been sufficient to separate GST dimers from potential HR1ab oligomers. Therefore, PRK1 HR1ab was concentrated and the protein loaded onto a glutathione sepharose column. This showed that there was some GST contamination that was removed (Figure 3.9A). The flow-through and wash were concentrated and loaded onto a Superdex 200 column. Two main peaks were observed (Figure 3.9B). Fractions were analysed by SDS-PAGE and C7-D6 were pooled and concentrated (Figure 3.9C) and the concentration was estimated by both measuring the absorbance as well as by a Bradford assay.

PRK3 HR1a and PRK3 HR1b were purified in the same way, with the initial glutathione sepharose column purifying the proteins from their main impurity, GST (Figures 3.10A and 3.11A). The flow-through and wash were concentrated in each case and loaded onto a Superdex 75 column with absorbance being monitored at 280 nm for PRK3 HR1a and 230 nm for PRK3 HR1b, which lacks aromatic amino acids (Figures 3.10B and 3.11B). For PRK3 HR1a, fractions C10-D2 were pooled and concentrated following SDS-PAGE analysis (Figure 3.10C) and the concentration was estimated by measuring the absorbance at 280 nm and with a Bradford assay. For PRK3 HR1b, fractions C11-D3 were pooled and concentrated following SDS-PAGE analysis (Figure 3.11C) and the concentration was estimated using a Bradford assay.

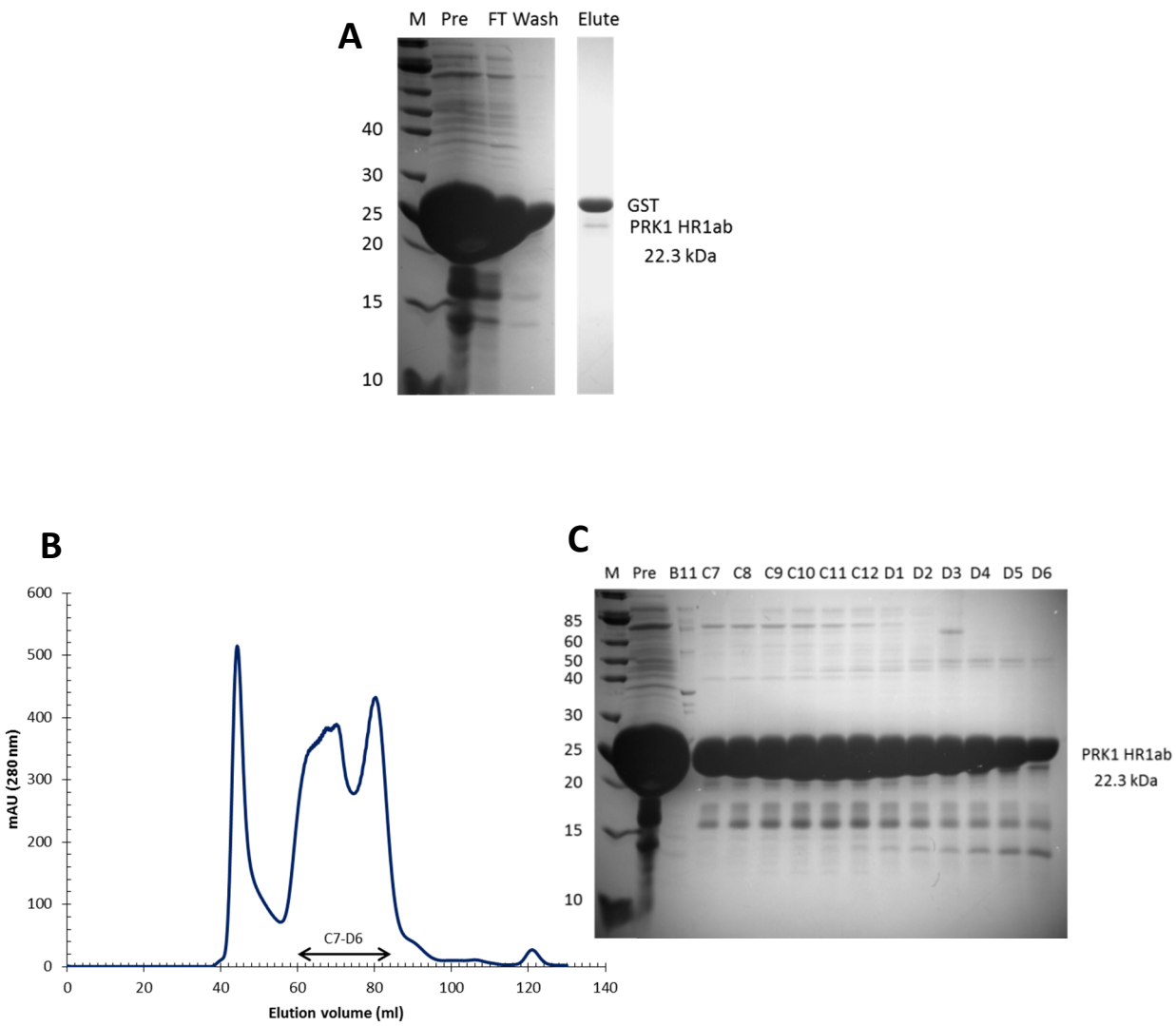


Figure 3.9 Further purification of PRK1 HR1ab and corresponding SDS-PAGE analysis. (**A**) SDS-PAGE analysis of purification on a glutathione sepharose column on a 15% Laemmli gel. (**B**) Gel filtration chromatogram of PRK1 HR1ab on an S200 column. Blue line indicates A_{280}. (**C**) SDS-PAGE analysis of chromatogram in (**B**) on a 15% Laemmli gel. 'Pre' for sample loaded on the column, FT for flow-through, Wash for wash with buffer A back to baseline and Elute for the column eluate after buffer B wash. The double arrows indicate the fractions pooled.

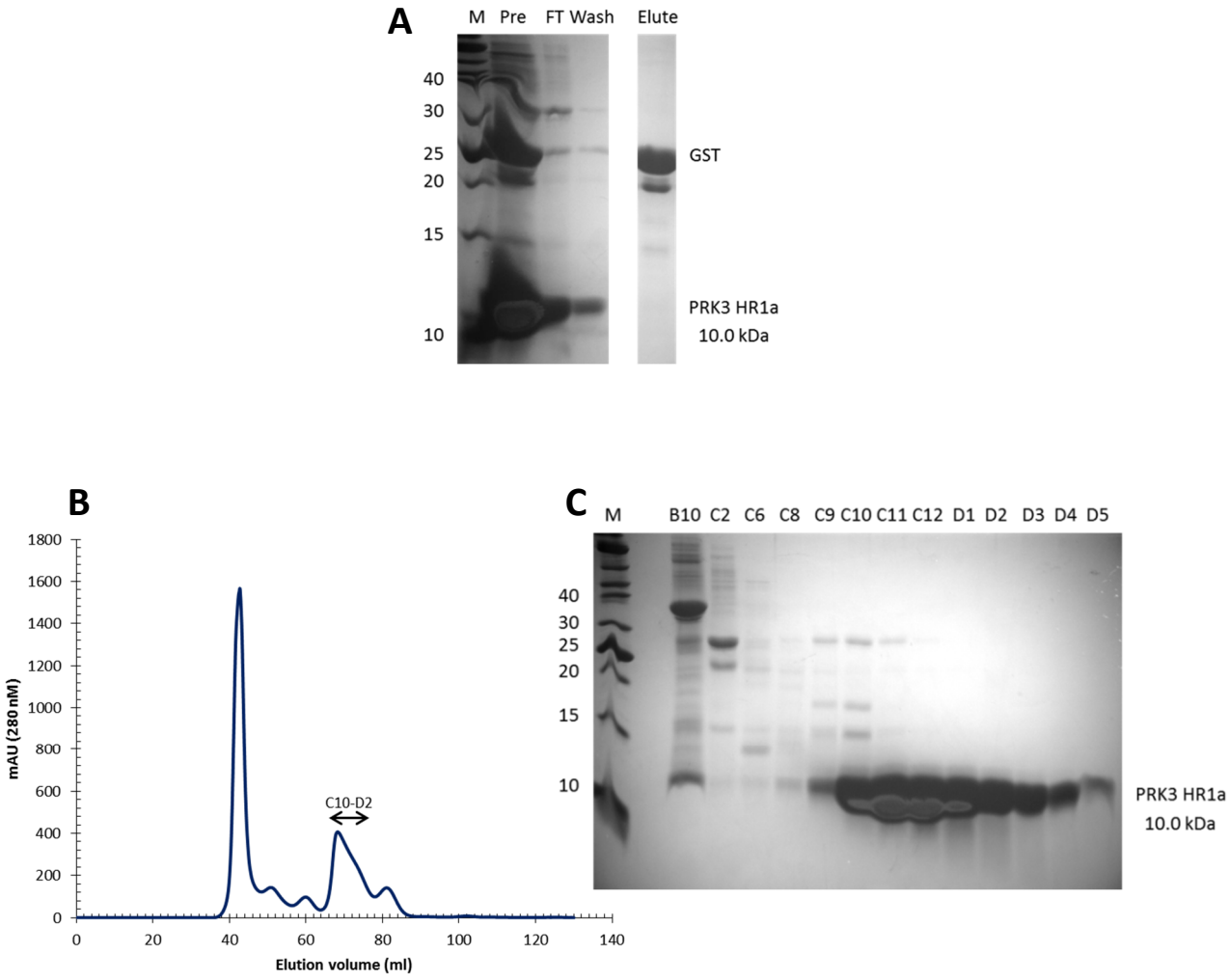


Figure 3.10 Further purification of PRK3 HR1a and corresponding SDS-PAGE analysis. (**A**) SDS-PAGE analysis of purification on a glutathione sepharose column on an 18% Laemmli gel. (**B**) Gel filtration chromatogram of PRK3 HR1a on an S75 column. Blue line indicates A_{280}. (**C**) SDS-PAGE analysis of chromatogram in (**B**) on an 18% Laemmli gel. 'Pre' for sample loaded on the column, FT stands for flow-through, Wash for wash with buffer A back to baseline and Elute for the column eluate after buffer B wash. The double arrows indicate the fractions pooled.

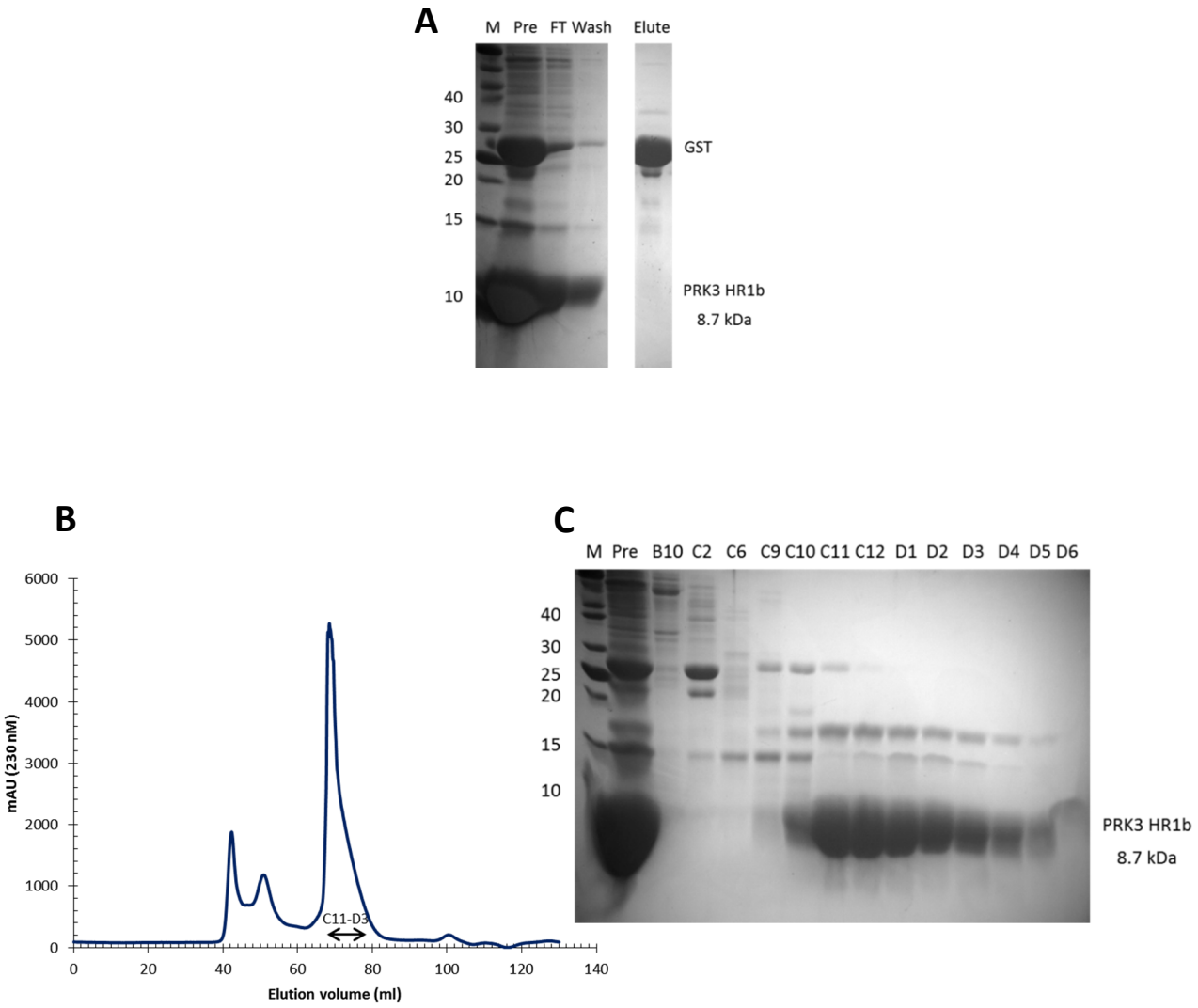


Figure 3.11 Further purification of PRK3 HR1b and corresponding SDS-PAGE analysis. (**A**) SDS-PAGE analysis of purification on a glutathione sepharose column on an 18% Laemmli gel. (**B**) Gel filtration chromatogram of PRK3 HR1b on an S75 column. Blue line indicates A_{230}. (**C**) SDS-PAGE analysis of chromatogram in (**B**) on an 18% Laemmli gel. 'Pre' for sample loaded on the column, FT for flow-through, Wash for wash with buffer A back to baseline and Elute for the column eluate after buffer B wash. The double arrows indicate the fractions pooled.

3.2.7 Further purification of PRK1 HR1a on a glutathione sepharose column followed by anion exchange chromatography and gel filtration

It was shown previously that loading the protein on a gel filtration column did not adequately purify the protein (Catherine Hutchinson, unpublished), suggesting that PRK1 HR1a may oligomerise. The protein was concentrated to 5 ml and loaded onto a glutathione sepharose column. This allowed purification of HR1a from the bulk of the main impurity, GST (Figure 3.12A). This purification step preceded the ion exchange step to increase the binding capacity of HR1a to the column (as GST is also known to bind). The flow-through and wash were pooled to a total of 33.5 ml. 6.7 ml aliquots were diluted to 50 ml with Buffer A1, reducing the [NaCl] to 20 mM. The protein was then loaded onto a ResourceQ column (Figure 3.12B) and the fractions analysed by SDS-PAGE (Figure 3.12C). Fractions B10-C4 were, concentrated, and loaded onto a Superdex 200 column. A single peak spanning 15 fractions, and corresponding to PRK1 HR1a, was observed (Figure 3.12D). Fractions C10-D3 and D7-D12 were pooled and concentrated (Figure 3.12E) and the protein concentration was estimated using absorbance at 280 nm and a Bradford assay.

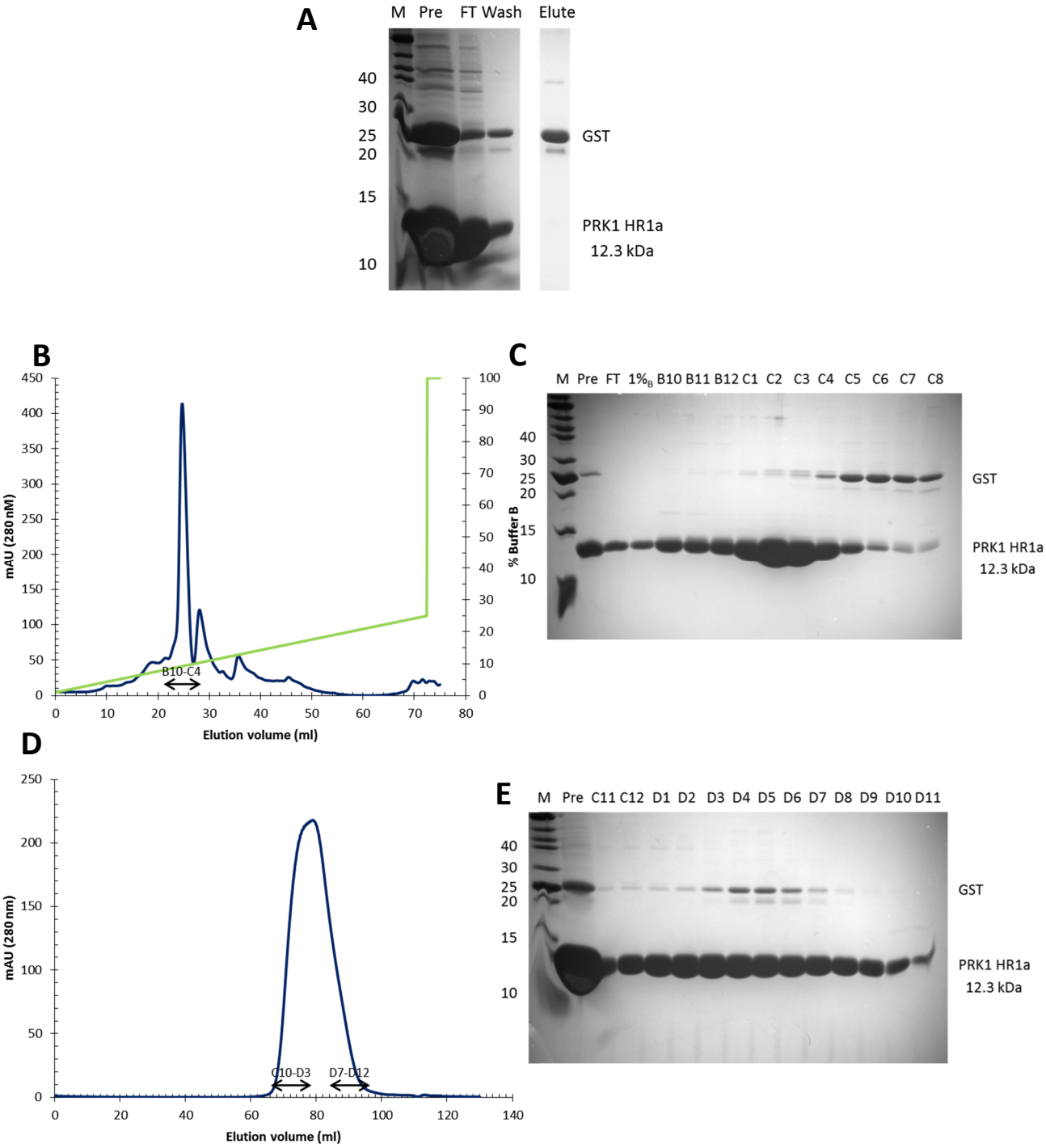


Figure 3.12 Further purification of PRK1 HR1a and corresponding SDS-PAGE analysis. (**A**) SDS-PAGE analysis of purification on a glutathione sepharose column on an 18% Laemmli gel. (**B**) Anion exchange chromatogram of PRK1 HR1a on a ResourceQ column. Blue line indicates A_{280}. Green line indicates increasing buffer B gradient. (**C**) SDS-PAGE analysis of chromatogram in (**B**) on an 18% Laemmli gel. (**D**) Gel filtration chromatogram of PRK1 HR1a on an S200 column. Blue line indicates A_{280}. (**E**) SDS-PAGE analysis of chromatogram in (**D**) on an 18% Laemmli gel. The double arrows indicate the fractions pooled. 'Pre' for sample loaded on the column, FT for flow-through, Wash for wash with buffer A back to baseline, Elute for the column eluate after buffer B wash and $1\%_B$ for wash with 1% buffer B back to baseline.

3.2.8 Further purification of PRK3 HR1ab and PRK3 HR1abc by cation exchange chromatograpy and gel filtration

PRK3 HR1ab has a theoretical isoelectric point of 11.0 and a predicted charge of +9.4 at pH 7.5. To purify the protein from GST and other higher molecular weight contaminants, ion exchange chromatography was employed. The protein was concentrated to 100 ml. 20 ml at a time were diluted up to 50 ml with Buffer A containing no NaCl, thereby diluting the [NaCl] to 200 mM. Some precipitation was observed at this stage, indicating that the protein is stable only in conditions of high ionic strength. The protein was then loaded onto a Heparin column (Figure 3.13A) and fractions were analysed by SDS-PAGE (Figure 3.13B). Fractions C7-D7 were pooled, concentrated to 2 ml and loaded onto a Superdex 200 column (Figure 3.13C). A single peak corresponding to PRK3 HR1ab was observed following SDS-PAGE analysis of the obtained fractions (Figure 3.13D). Fractions D3-D10 were pooled, concentrated and the concentration was estimated using a Bradford assay.

The purification of PRK3 HR1abc had to undergo extensive optimisation in order to obtain a protein sample of high yield and purity (both from contaminants such as GST and also from nucleic acids). PRK3 HR1abc is a highly positively charged protein, with a theoretical isoelectric point of 10.8 and predicted charge of +12 at pH 7.5. Previous attempts at purifying a His-tagged version of this protein showed that it is impossible to sufficiently purify the protein solely by gel filtration. In addition, absorbance at 260 nm suggested that there was nucleic acid contamination. Ion exchange chromatography was therefore employed as a way of removing nucleic acid contamination, which is crucial for identifying the accurate molecular weight of different species in analytical ultracentrifugation experiments, and also purifying the protein from contaminating species such as GST. The conditions for loading onto the ion exchange column were optimised from several large scale expressions and were found to be 0.22 mg/ml protein, 200 mM NaCl in 50 ml. This achieved the following: (a) 200 mM NaCl was sufficiently low for the protein to stick to the heparin ligand, (b) diluting the NaCl led to some protein precipitation, which was minimised when the final protein concentration was 0.22 mg/ml or lower (c) the Heparin column was not saturated, i.e. maximising usage of the column capacity and minimising the number of separate aliquots that needed to be purified on the column. Collectively, this meant that the

protein was purified from several, usually around 6, separate method runs. Specifically, PRK3 HR1abc washes were concentrated to 120 ml. 20 ml aliquots were diluted to 50 ml with Buffer A2, reducing the [NaCl] to 200 mM. The protein was then loaded onto a Heparin column (Figure 3.14A). Fractions were analysed by SDS-PAGE and C8-D12 were pooled (Figure 3.14B). This was done six times in total, once per 20 ml aliquot. The protein was then concentrated to 2 ml and loaded onto a Superdex 200 column (Figure 3.14C). Two peaks were observed, indicating that this tri-domain is also oligomerising. Fractions were analysed by SDS-PAGE and C7-D5 were pooled and concentrated (Figure 3.14D). The concentration was estimated by a Bradford assay. The protein is stable at 4 ^{o}C but precipitates as it warms up. Therefore care was taken to ensure the protein was under cold conditions during all stages of purification. Interestingly, the precipitation appeared to be reversible; placing the protein back on ice caused the protein to go back into solution in a few seconds. An analogous phenomenon was observed with PRK3 HR1ab.

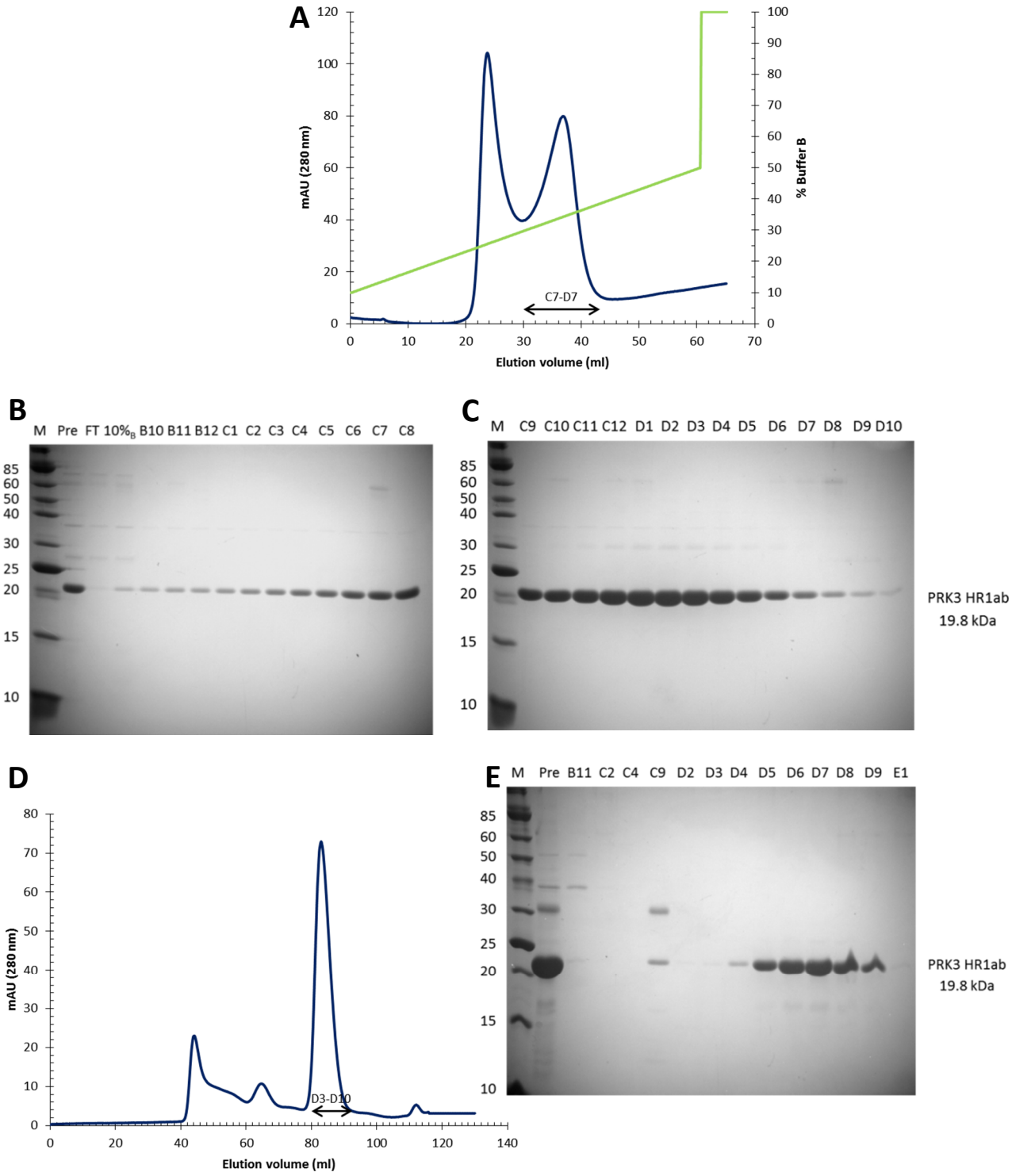


Figure 3.13 Further purification of PRK3 HR1ab and corresponding SDS-PAGE analysis. (**A**) Cation exchange chromatogram of PRK3 HR1ab on a Heparin column. Blue line indicates A_{280}. Green line indicates increasing buffer B gradient. (**B**) SDS-PAGE analysis of chromatogram in (**A**) on a 15% Laemmli gel. (**C**) Gel filtration chromatogram of PRK3 HR1ab on an S75 column. Blue line indicates A_{280}. (**D**) SDS-PAGE analysis of chromatogram in (**C**) on a 15% Laemmli gel. The double arrows indicate the fractions pooled. 'Pre' for sample loaded on the column, FT for flow-through and $10\%_B$ for wash with 10% buffer B back to baseline.

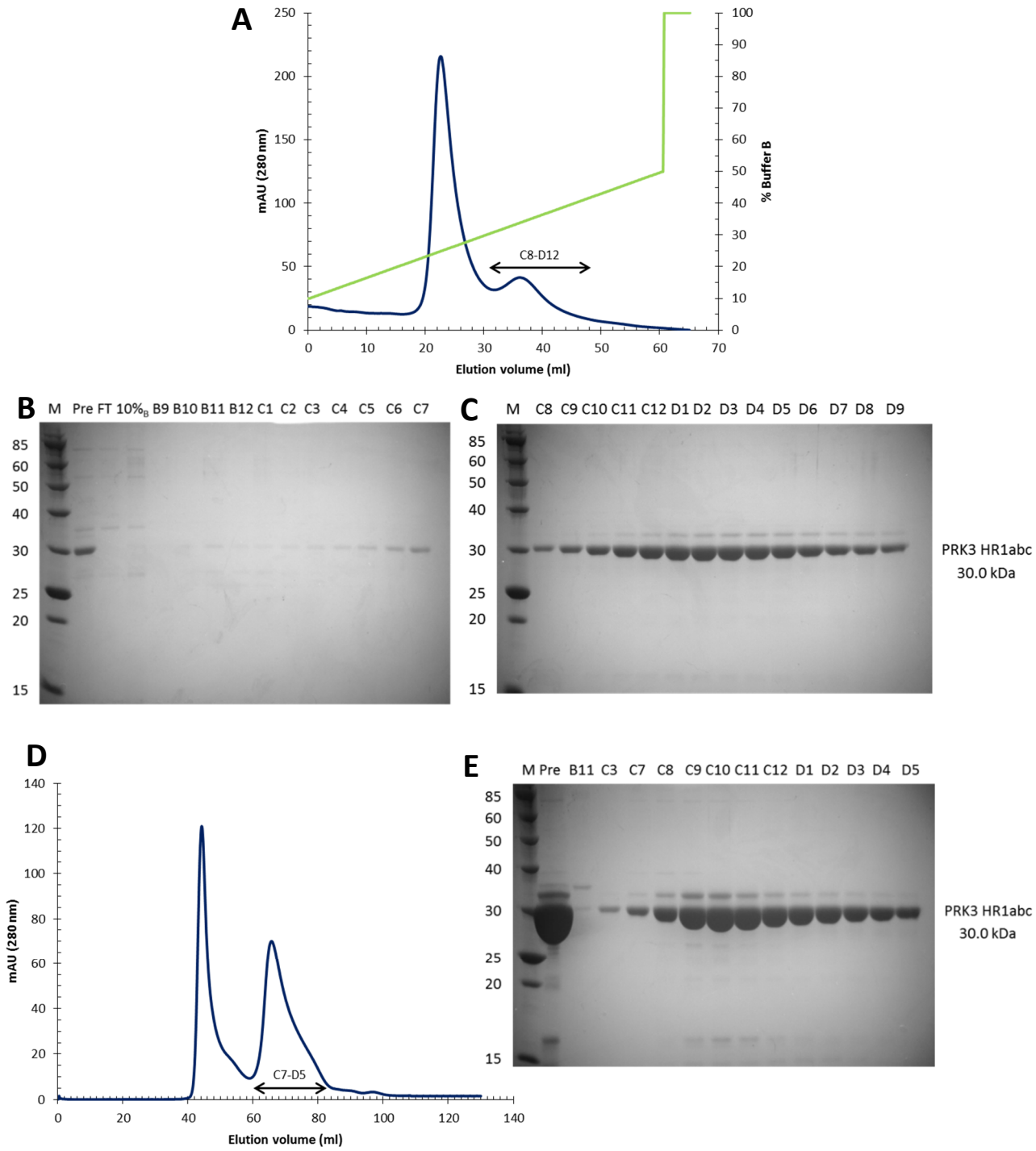


Figure 3.14 Further purification of PRK3 HR1abc and corresponding SDS-PAGE analysis. (**A**) Cation exchange chromatogram of PRK3 HR1ab on a Heparin column. Blue line indicates A_{280}. Green line indicates increasing buffer B gradient. (**B**) SDS-PAGE analysis of chromatogram in (**A**) on a 12% Laemmli gel. (**C**) Gel filtration chromatogram of PRK3 HR1abc on an S75 column. Blue line indicates A_{280}. (**D**) SDS-PAGE analysis of chromatogram in (**C**) on a 12% Laemmli gel. The double arrows indicate the fractions pooled. 'Pre' for sample loaded on the column, FT for flow-through and $10\%_B$ for wash with 10% buffer B back to baseline.

3.2.9 Yields of purified proteins

The accurate concentration of each protein was determined and purity to 95% was checked by amino acid analysis. A correction factor, which translates a protein concentration estimated by absorbance at 280 nm or with a Bradford assay to the true concentration, was used in subsequent purifications of each protein. The yields of each of the purified proteins are summarised in table 3.2.

PRK1 HR1c and PRK3 HR1c, which had not been previously purified, were confirmed by MALDI Mass Spectrometry to be intact and not breaking down at the end of the purification procedure (see Appendix 2).

	Protein	**Volume of *E. coli* culture (L)**	**Volume of purified protein (ml)**	**Concentration (mM)**
PRK1	HR1a	4	1	1.7
	HR1b	4	0.675	6.5
	HR1c	4	1	3.02
	HR1ab	4	1.5	1.2
	HR1abc	6	1	0.65
PRK3	HR1a	4	1	0.52
	HR1b	4	1	0.45
	HR1c	4	1	0.800
	HR1ab	4	1	0.057
	HR1abc	6	1	0.25

Table 3.2 Final yields of purified proteins obtained. Amino acid analysis was used to accurately determine the protein concentration.

3.3 Investigating the secondary structure of HR1 domains by circular dichroism

3.3.1 Introduction

To determine whether the purified HR1 proteins are structured, their secondary structures were investigated by circular dichroism. The predicted α-helical structure of these domains makes them ideal candidates for circular dichroism experiments. In addition, the thermal stability of the HR1c domains was studied to allow comparison with the published data on the HR1a and HR1b constructs (Hutchinson *et al.*, 2013).

3.3.2 Theory

Circular dichroism (CD) is an excellent method for obtaining an estimate of the total amount of secondary structure present in a population of protein molecules in solution. This includes α-helices, β-sheets, polyproline helices and random coils (Greenfield, 2006). CD is defined as the differential absorption of left-handed and right-handed circularly polarised light. Chiral molecules can absorb left-handed and right-handed circularly polarised light to a different extent and this can give rise to a CD signal.

The oscillating electric field of a beam of polarised light can be described in terms of two vectors which rotate circularly, one clockwise to the left (L) and the other counterclockwise to the right (R). If these two components are absorbed equally or not absorbed at all by the sample, then their recombination will regenerate the radiation polarised in the starting plane (Kelly *et al.*, 2005). If L and R have different magnitudes, as a result of being absorbed to different extents by the sample, then elliptical polarisation is generated, which gives rise to the CD signal. This difference in absorbance, ΔA, between the L and R circularly polarised components is the quantity measured by CD spectropolarimeters (Figure 3.15A).

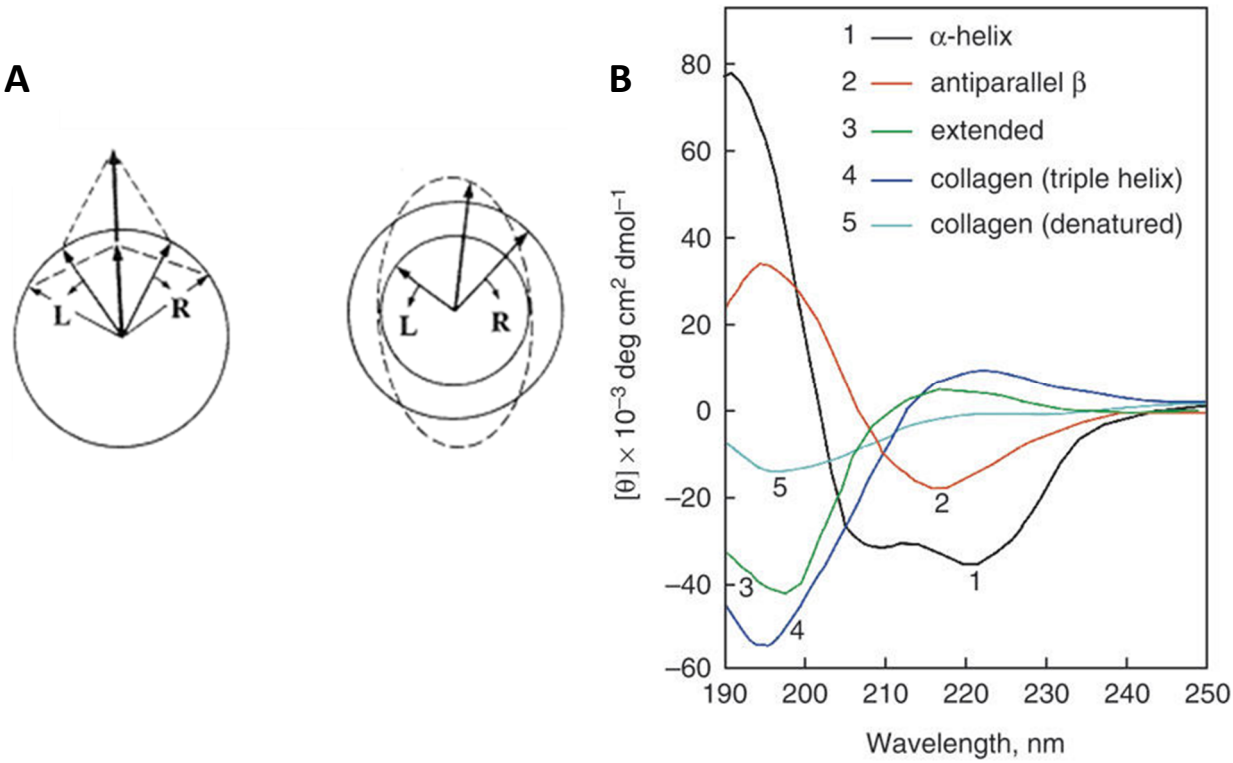


Figure 3.15 Circular dichroism principles. (**A**) The CD effect. Left (L) and right (R) circularly polarised components combine to generate plane polarised radiation. Left panel: The magnitude of the two components is the same. The resulting plane polarised radiation is not elliptically polarised. Right panel: The magnitude of the two components is different resulting in elliptically (dashed line) plane polarised radiation. Figure adapted from Kelly *et al.* (2005). (**B**) Overlaid representative CD spectra of proteins with different secondary structures. Figure adapted from Greenfield (2006).

$$\Delta A = A_L - A_R \qquad 3.1$$

CD spectra are often expressed in ellipticity

$$\theta = tan^{-1}\left(\frac{b}{a}\right) \qquad 3.2$$

where θ is the ellipticity in degrees and b and a are the minor and major axes, respectively, of the ellipse arising from the L and R components.

This is directly proportional to ΔA.

$$\theta = 32.98\Delta A \qquad 3.3$$

In the common CD experiment, the ellipticity is recorded during a wavelength scan in the far UV region of proteins, e.g. 185 nm to 240 nm. The chromophore in proteins is the peptide bond. At a wavelength of 210-220 nm, there is a π → π* energy transition from free

electrons in carbonyl oxygens. In addition, there is a stronger $\pi \rightarrow \pi^*$ energy transition involving the carbonyl π-electrons. These transitions depend on the peptide ϕ and ψ torsion angles and therefore differ between the different secondary structures that can be present in proteins (Kelly *et al.*, 2005) (Figure 3.15B). In particular, in the case of an α-helical protein, the $\pi \rightarrow \pi^*$ transitions lead to a positive signal at 190 nm and negative signals at 208 nm and 222 nm (Kelly *et al.*, 2005; Ranjbar & Gill, 2009).

Online tools such as DICHROWEB exist, which facilitate the analysis of recorded CD spectra and fit the data according to reference datasets of proteins with specific secondary structures, e.g. fully α-helical. This analysis can lead to an estimate of the total secondary structure in solution, which can be informative of the secondary structure of the molecule, assuming a single population of species. This has additional advantages in studies which investigate secondary structure changes upon ligand binding.

The thermal stability of proteins can also be assessed by CD (Ranjbar & Gill, 2009). A wavelength at which a particular secondary structure feature of the protein leads to a strong signal is chosen, e.g. 222 nm for α-helical proteins. Then the ellipticity at that wavelength is monitored with increasing temperature, giving rise to a sigmoidal curve, the mid-point of which is the melting temperature, T_m. This use of CD has many applications, for example in the screening of the stability of proteins in conditions of different buffer, pH, ionic strength of the solution etc. (Kelly *et al.*, 2005).

3.3.3 Results

All HR1 constructs purified contain secondary structure (Figures 3.16-3.17) and have a % helicity ranging between 44-94%. For most HR1 domains, % helicity is similar to that predicted by the jpred server. Any considerable differences between predicted and recorded helicities are likely to be due to the high rmsd of the particular data reconstruction on DICHROWEB.

The two HR1c domains were more structured than predicted, possibly because they form additional helical structure which is not predicted by jpred.

PRK3 HR1ab was 89% helical and more helical than either PRK3 HR1a or PRK3 HR1b were as separate domains, suggesting an interaction between these two domains. Surprisingly, PRK3 HR1ab in this study had more than double the helicity than His-tagged PRK3 HR1ab studied in Hutchinson *et al.* (2013), suggesting that the fusion tag prevents the correct folding of the PRK3 di-domain.

The two HR1abc proteins were slightly more helical than predicted, suggesting that the three HR1 domains are not independent from each other and may form additional secondary structure.

The results are summarised in Table 3.3.

Construct	jpred	Structure	CD			
			Helix 1	Helix 2	Total	rmsd
PRK1 HR1a	0.536	0.652	0.38	0.22	0.60	0.014
PRK1 HR1b	0.759	0.795	0.49	0.27	0.76	0.010
PRK1 HR1c	0.701		0.61	0.33	0.94	0.008
PRK1 HR1ab	0.605		0.27	0.17	0.44	0.019
PRK1 HR1abc	0.607		0.57	0.24	0.81	0.007
PRK3 HR1a	0.465		0.29	0.19	0.48	0.019
PRK3 HR1b	0.772		0.50	0.25	0.75	0.012
PRK3 HR1c	0.725		0.59	0.29	0.88	0.009
PRK3 HR1ab	0.661		0.58	0.31	0.89	0.008
PRK3 HR1abc	0.639		0.49	0.29	0.78	0.012

Table 3.3 Summary of CD data. The determined helicity from the fit data on DICHROWEB was compared to the predicted helicity from jpred and also the actual helical content from known structures in the case of PRK1 HR1a and PRK1 HR1b.

The thermal stability of HR1c domains has never been investigated. To test the thermal stability of the HR1c domains from PRK1 and PRK3, thermal melts were recorded by monitoring ellipticity at 222 nm with increasing temperature (Figure 3.18). The T_m of PRK1 HR1c domain is approximately 53 ^{o}C, while that of PRK3 HR1c is approximately 58 ^{o}C. Interestingly, while the T_m of PRK3 HR1c is higher than that of its PRK1 counterpart,

secondary structure loss was observed earlier and was only complete at 80 ^{o}C. Contrastingly, the melting curve of PRK1 HR1c is much sharper, with most of the secondary structure lost by 65 ^{o}C. This is indicative of helical elements in PRK3 HR1c which differ significantly in their stability, whereas they are more uniform in the case of PRK1 HR1c. It's possible that any differences are due to different oligomeric states of the PRK1 and PRK3 HR1c domains, as suggested by gel filtration (see section 3.2.5). The broad temperature range over which PRK3 HR1c secondary structure was lost is unlikely to be due to a mixed population of structured and unstructured molecules because the wavelength scan on this protein indicated a very high percentage helicity. Furthermore, when the proteins were cooled down back to 20 ^{o}C, the secondary structure was almost completely restored in the case of both proteins. Both HR1c domains have a predicted C-terminal helix which is shorter than the main two helices of the domain and is exclusive to the HR1c domains. As only a single T_m was apparent from the data, this would suggest that this additional helix may not be independent of the other two helices.

The T_m values of the HR1c domains are higher than those of PRK1 HR1a, PRK2 HR1b, PRK3 HR1a and PRK3 HR1b and very similar to those of PRK1 HR1b and PRK2 HR1a (Hutchinson *et al.*, 2013). The existence of two broad groups of HR1 domains, those with higher and those with lower melting temperatures may be key for how these domains differentially interact with GTPases.

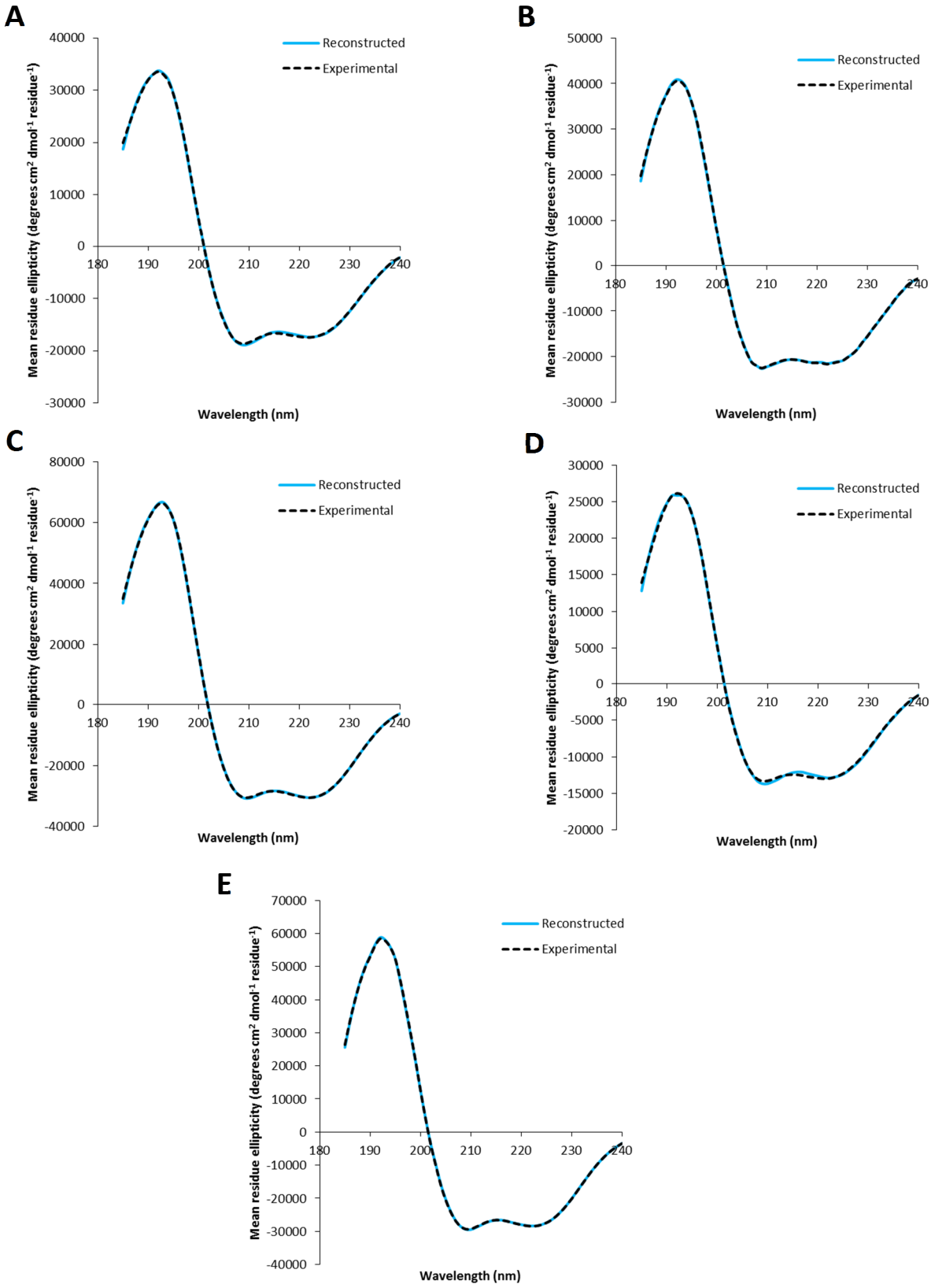


Figure 3.16 PRK1 construct circular dichroism wavelength scans. (**A**)-(**E**) The scans were recorded with 0.2 mg/ml protein. The experimental data is overlaid with the reconstructed data from the DICHROWEB analysis. (**A**) PRK1 HR1a (**B**) PRK1 HR1b (**C**) PRK1 HR1c (**D**) PRK1 HR1ab (**E**) PRK1 HR1abc

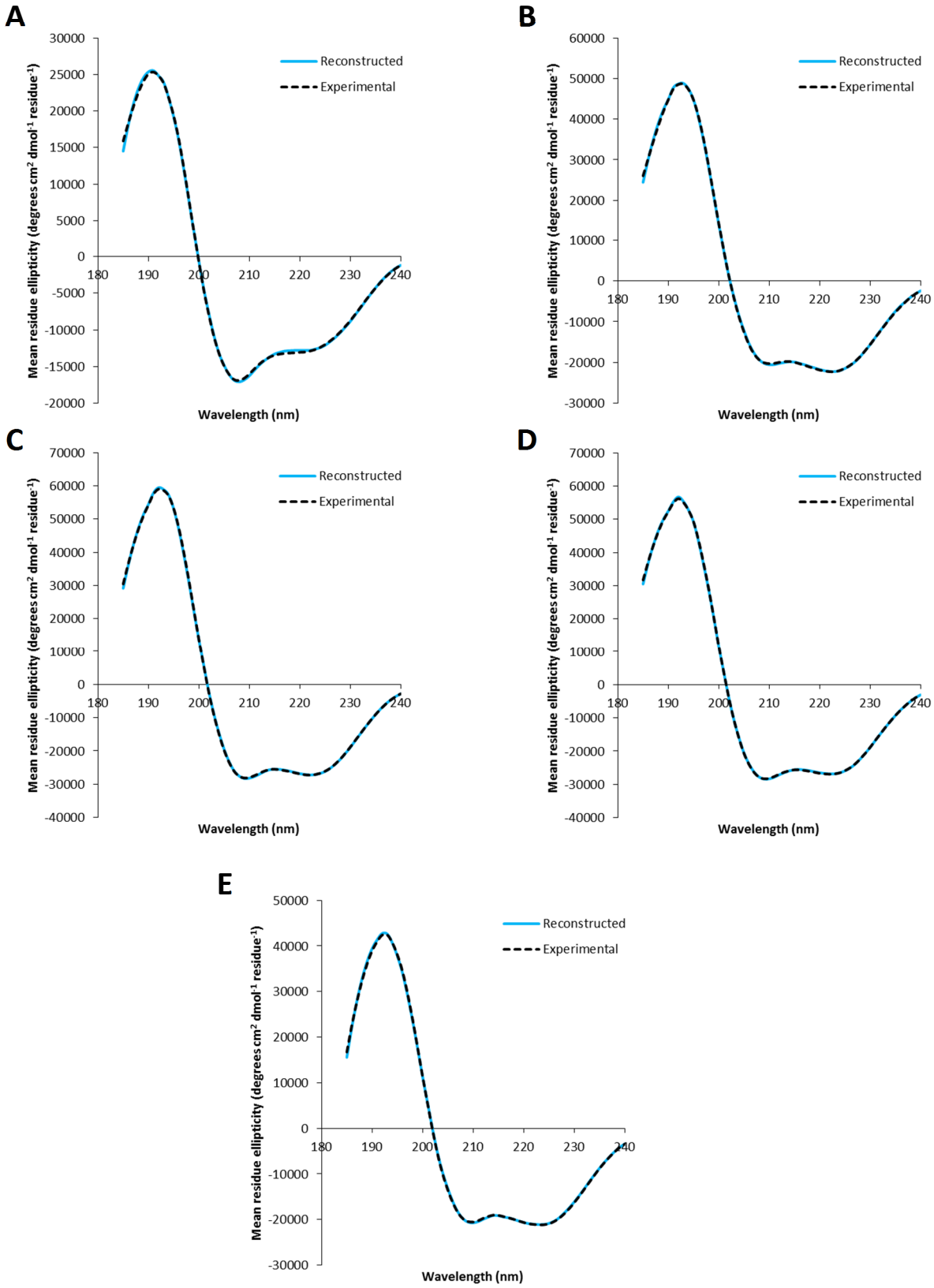


Figure 3.17 PRK3 construct circular dichroism wavelength scans. (**A**)-(**E**) The scans were recorded with 0.2 mg/ml protein. The experimental data is overlaid with the reconstructed data from the DICHROWEB analysis. (**A**) PRK3 HR1a (**B**) PRK3 HR1b (**C**) PRK3 HR1c (**D**) PRK3 HR1ab (**E**) PRK3 HR1abc

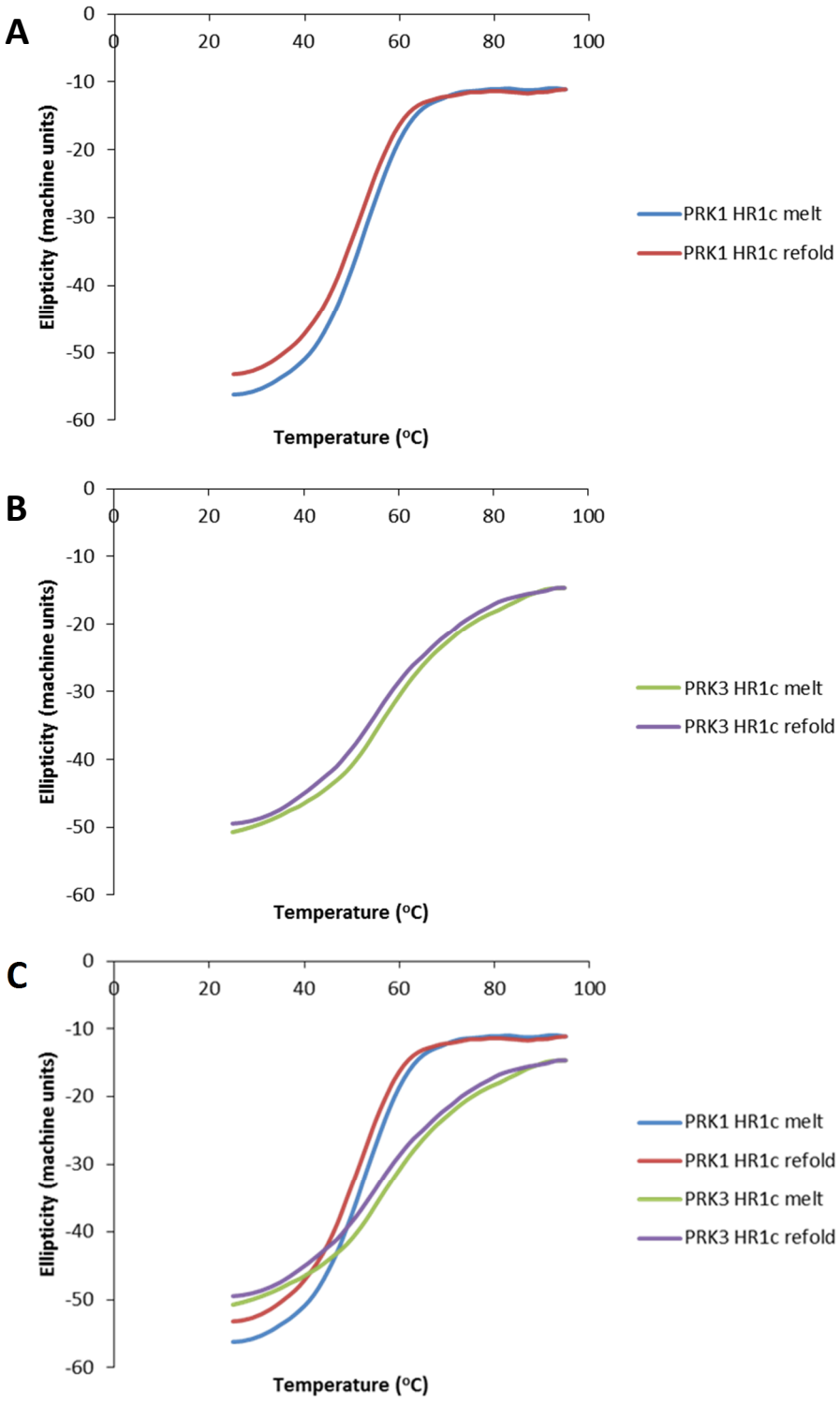


Figure 3.18 Thermal stability comparison between PRK1 HR1c and PRK3 HR1c. (**A**)-(**C**) The thermal melt scans was recorded by recording the ellipticity at 222 nm with increasing temperature with 0.2 mg/ml protein. Both the melt and the cooling down (refold) are shown. (**A**) PRK1 HR1c melt and refold (**B**) PRK3 HR1c melt and refold (**C**) Overlay

3.4 Investigating the oligomeric state of HR1 domains by analytical ultracentrifugation

3.4.1 Theory

Interacting proteins often exist in reversible equilibria between free and bound states. Some proteins can reversibly self-associate and this can be summarised with a system in equilibrium:

$$A + A \rightleftharpoons A_2 \quad 3.4$$

The association rate of the reaction can be described as

$$Rate = k_{on}[A]_{free}[A]_{free} \quad 3.5$$

$$Rate = k_{on}[A]_{free}{}^2 \quad 3.6$$

where k_{on} is the association rate constant and $[A]_{free}$ is the concentration of monomeric molecule A.

Similarly, we can write an expression for the dissociation rate

$$Rate = k_{off}[A]_{complex} \quad 3.7$$

where k_{off} is the dissociation rate constant and $[A]_{complex}$ is the concentration of the complex.

At equilibrium, the association and dissociation rates are equal such that

$$k_{on}[A]_{free}{}^2 = k_{off}[A]_{complex} \quad 3.8$$

The equilibrium constant, K_{eq}, which describes a self-associating system in equilibrium, can be defined as

$$K_{eq} = \frac{k_{on}}{k_{off}} = \frac{[A]_{complex}}{[A]_{free}{}^2} \quad 3.9$$

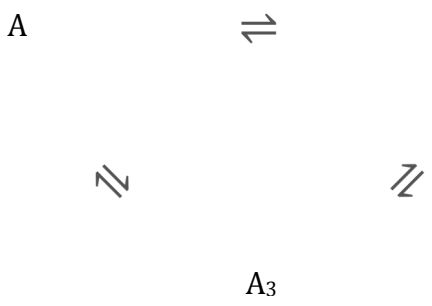
The reciprocal of this is the dissociation constant, K_d, which is used to describe macromolecular interactions.

$$K_d = \frac{k_{off}}{k_{on}} = \frac{[A]_{free}{}^2}{[A]_{complex}} \qquad 3.10$$

The total protein concentration can also be defined as

$$[A]_{total} = [A]_{free} + 2[A]_{complex} \qquad 3.11$$

Very often systems exist in reversible equilibria which involve more than two species. These are called linked equilibria (reviewed in Pollard, 2010) and the linked triangle below represents possible interactions between all species in a monomer-dimer-trimer system.



It is important to note that each of these reactions can be described by a separate dissociation constant, and in turn different association and dissociation rates for the forward and reverse reactions, respectively. The complexity of characterising such a system becomes exponentially difficult with an increasing number of different species in solution. However, the first step when characterising the oligomeric state of a macromolecule is to determine the number of species in solution and their identity. This should ideally be done at a range of concentrations, as with increasing concentration the proportion of higher molecular weight species should increase.

In this work, the oligomeric state of the HR1 domains was investigated by analytical ultracentrifugation (AUC). In contrast to analytical gel filtration and native SDS-PAGE, AUC can accurately determine the molecular weight of the species in solution and also provide information on their shape (Balbo & Schuck, 2005). In the 1990s analytical ultracentrifugation re-emerged as a powerful method of studying proteins and is considered

to be the "gold standard" in analysing self-association. It is a first principles technique that studies molecules in solution and therefore it has an advantage over other techniques (e.g. bio-layer interferometry) in that molecules are not tethered on surfaces. AUC can give an estimate of the amount of each species present in solution and its molecular weight. It is a powerful method of studying heterogeneous protein interactions and it is also extremely sensitive to sample purity which makes it very useful in industry.

In AUC, a sample and reference are loaded in two separate sectors of a cell (Figure 3.19A). The cell is then loaded in a rotor, which is spun at high speeds and the sedimentation process which occurs is monitored by an optics component that is present in the ultracentrifuge (Figure 3.19B). Centrifugation causes sedimenting species to yield sigmoidal concentration profiles whose inflexion points are referred to as sedimenting "boundaries" (Figure 3.19C). For solutions containing two or more species of adequately different size, an equal number of boundaries may be observed.

A sedimenting particle (in this case a protein with its solvation shell), of mass *m*, experiences three forces (Figure 3.19D) (Ralston, 1993). Firstly, it experiences the gravitational force, F_s, which is equal to the mass times the centrifugal acceleration, $\omega^2 r$:

$$F_s = m\omega^2 r = \frac{M}{N_A}\omega^2 r \qquad 3.12$$

where *m* is the mass of the particle in grams, ω is the angular velocity in radians per second, *r* is the distance of the particle from the rotation axis, *M* is the solute molar weight in g/mol and N_A is Avogadro's number.

Secondly, it experiences a buoyant force, F_b. According to Archimedes' principle,

$$F_b = -m_0\omega^2 r \qquad 3.13$$

where m_0 is the mass of the displaced solvent.

From Archimedes' principle,

$$m_0 = m\bar{v}\rho = \frac{M}{N_A}\bar{v}\rho \qquad 3.14$$

where $\bar{v}$ is the partial specific volume in ml, and ρ is the solvent density in g/ml.

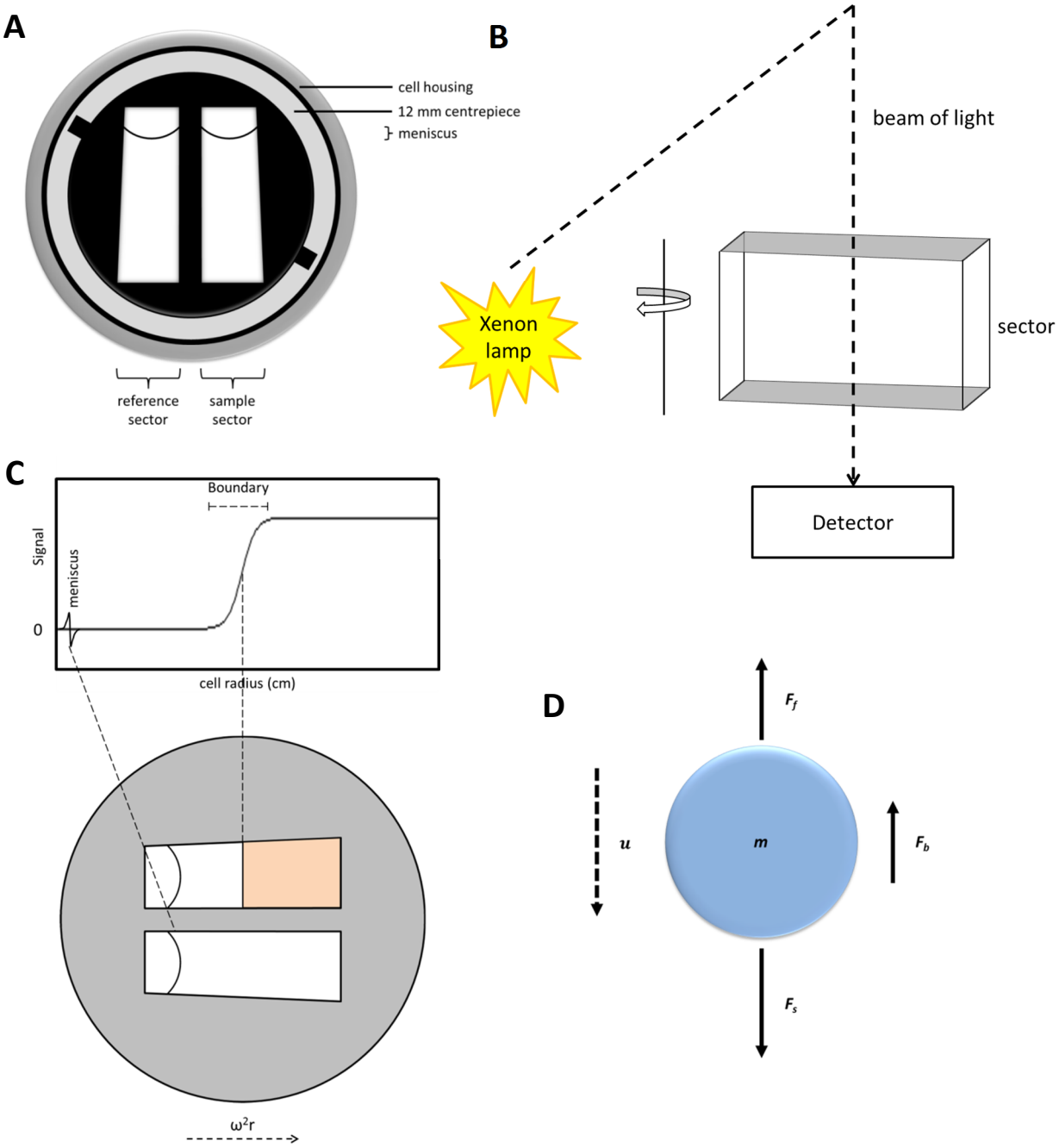


Figure 3.19 Analytical ultracentrifugation setup and principles. (**A**) Schematic of an AUC cell. The reference and sample sectors are equally filled with matching buffer and protein, respectively. (**B**) A xenon flash lamp creates a beam of light which passes through the sectors of an AUC cell placed in a revolving rotor and terminates at the detector. (**C**) Sedimentation gives rise to a sigmoidal boundary which can be analysed to derive the sedimentation coefficient, *S*, and the diffusion coefficient, D. Figure adapted from Ralston (1993). (**D**) Forces acting upon a sedimenting particle. These are the gravitational force, F_s, the buoyant force, F_b and the frictional force, F_f. *u* stands for the sedimenting velocity of the particle.

Substituting (3.14) in (3.13) gives

$$F_b = -\frac{M}{N_A}\bar{v}\rho\omega^2 r \qquad 3.15$$

Thirdly, the sedimenting particle experiences the frictional force, F_f, which is equal to the frictional coefficient, *f*, times the sedimenting velocity of the particle, u.

$$F_f = -fu \qquad 3.16$$

Shortly after the beginning of an experiment the three forces come into equilibrium, such that

$$Fs + Fb + Ff = 0 \qquad 3.17$$

Substituting (3.12), (3.15) and (3.16) in (3.17) gives

$$\frac{M}{N_A}\omega^2 r - \frac{M}{N_A}\bar{v}\rho\omega^2 r - fu = 0 \qquad 3.18$$

Collecting the common terms results in

$$\frac{M}{N_A}\omega^2 r(1 - \bar{v}\rho) - fu = 0 \qquad 3.19$$

Rearranging gives

$$\frac{M(1 - \bar{v}\rho)}{N_A f} = \frac{u}{\omega^2 r} = s \qquad 3.20$$

$$s = \frac{velocity}{acceleration} = \frac{u}{\omega^2 r} \qquad 3.21$$

where *s* is the sedimentation coefficient and is equal to the velocity divided by the acceleration (which is the centrifugal field strength, $\omega^2 r$).

The sedimentation coefficient can be directly related to a sedimenting boundary

$$s = \frac{u}{\omega^2 r} = \frac{\frac{dr_{boundary}}{dt}}{\omega^2 r} \qquad 3.22$$

and it is the first quantity directly measurable from the data.

Furthermore, it can be deduced by integration of (3.22) that

$$s\omega^2 t = ln\left(\frac{r_{boundary}}{r_m}\right) \qquad 3.23$$

where r_m is the radial position of the meniscus in the cell.

The sedimentation coefficient depends on the solvent density, viscosity and also on the partial specific volume of the solute. It is often useful, for comparative purposes, to convert the *s*-value from a particular experiment to the corresponding *s*-value of the species in standard conditions, where the solvent is water and the temperature is 20 ^{o}C (Lebowitz *et al.*, 2002). The sedimentation coefficient in these standard conditions, $s_{20,w}$, can be calculated as follows

$$s_{20,w} = \frac{(1-\bar{v}\rho)_{20,w}}{(1-\bar{v}\rho)_{4,b}}\frac{\eta_{20,w}}{\eta_{4,b}}s_{4,b} \qquad 3.24$$

where η stands for the solvent viscosity, *w* stands for water, *b* for the protein buffer and 4 for 4 ^{o}C.

The second measurable quantity from the data is the diffusion coefficient, *D*. This can be measured from the spreading of a sedimenting boundary. Einstein's rule relates *D* to the frictional coefficient, *f*. The two quantities are inversely proportional.

$$D = \frac{RT}{N_A f} \qquad 3.25$$

where *R* is the gas constant and *T* is the absolute temperature in Kelvin.

Rearranging this gives an expression for the frictional coefficient

$$f = \frac{RT}{N_A D} = \frac{kT}{D} \qquad 3.26$$

where *k* is the Boltzmann constant, equal to 1.38 x 10^{-23} J K^{-1}.

Substituting in equation (3.20) gives

$$s = \frac{u}{\omega^2 r} = \frac{M(1-\bar{v}r)}{N_A f} = \frac{MD(1-\bar{v}r)}{RT} \quad 3.27$$

Dividing both sides by *D* gives

$$\frac{s}{D} = \frac{M(1-\bar{v}r)}{RT} \quad 3.28$$

which is known as the Svedberg equation that describes a system made up of two components (solute and solvent). This equation relates the sedimentation coefficient, the diffusion coefficient and the buoyant mass of the particle, $M(1-\bar{v}r)$, and is the basis of the analysis of sedimentation profiles.

Sedimentation is also dependent on sample viscosity. In particular, the Stokes equation relates the frictional coefficient, f, to the viscosity, η, of the sample and the hydrodynamic radius of the solute (also known as Stokes radius, R_0), which is related to the asymmetry and solvation of the molecule.

$$f = 6\pi\eta R_0 \quad 3.29$$

Substituting this in equation (3.20) gives

$$s = \frac{u}{\omega^2 r} = \frac{M(1-\bar{v}r)}{N_A 6\pi\eta R_0} \quad 3.30$$

When data are analysed, the fitted frictional coefficient f is divided by the frictional coefficient of a hydrodynamically equivalent sphere, f_0, which yields the frictional ratio, $\frac{f}{f_0}$. This is a measure of the shape of the molecule. A more elongated molecule, with a higher hydrodynamic radius, will therefore also result in a higher frictional coefficient and, in turn, in a higher frictional ratio. Globular proteins tend to have frictional ratios 1-1.2 and more elongated proteins tend to have frictional ratios higher than 1.2.

Lamm was the first to describe the sedimentation of a boundary with a mathematical term. The Lamm equation (Lamm, 1929) describes the evolution of the concentration distribution, with respect to time, of an ideally sedimenting species

$$\frac{\partial c}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[rD\frac{\partial c}{\partial r} - sw^2r^2c\right] \quad 3.31$$

where c is the concentration, D is the diffusion coefficient, s is the sedimentation coefficient, r is the radius from the centre of rotation, ω is the rotor angular velocity and t is time.

Solving this equation can determine *s*, *D* and *M* according to the Svedberg equation. Programs such as SEDFIT provide numerical solutions to the Lamm equation and thus the rapid modelling of non-interacting species based on the recorded SV data. This method can account for sample heterogeneity and yields a differential sedimentation coefficient distribution, *c(s)*

$$a(r,t) \cong \int_{s_{min}}^{s_{max}} c(s)\chi(s,F,r,t)ds \qquad 3.32$$

where $a(r,t)$ is the acquired experimental data, $c(s)$ is the continuous distribution of s-values (and is proportional to protein concentration) and $\chi(s,F,r,t)$ is a Lamm Equation solution which is dependent on $D(s)$.

In the $c(s)$ model, the data are fit by assuming a single frictional ratio represents all different species in solution. While in certain cases this may have no significant impact on data interpretation, e.g. in a mixture of globular proteins with frictional ratios ranging from 1.2-1.4, this could be problematic if macromolecules with significantly different frictional ratios are present in solution. From equation (3.20), if the frictional ratio is, for example, underestimated, then the molecular weight corresponding to a measured s-value will be underestimated too. It is possible, if further information is available on the species present, to fit the data using a $c\left(s,\frac{f}{f_0}\right)$ model, where both the *s*-value and the frictional ratio are varied. This method can be used to provide a better understanding of the heterogeneity of a solution and can be aided by prior knowledge of species present and their possible frictional ratios. The $c(s,\frac{f}{f_0})$ distribution may also be used to examine protein conformational changes, such as the apoE3 protein having a more extended conformation in a submicellar environment (Chou *et al.*, 2011). However, it is subject to problems, primarily the frequent lack of prior knowledge on the shape of the various species. Integration of the *c(s)* distribution of *s*-values gives the abundance of each species in solution.

For non-interacting species or for species in slow chemical exchange, the true s-values of each of the sedimenting species are measured. The $c(s)$ model however does not account for species which are reversibly interacting on the time-scale of the experiment, i.e. in intermediate-to-fast chemical exchange. For fast exchanging equilibria, e.g. monomer-dimer equilibrium, a single sedimenting boundary may be observed. This gives rise to a weighted-average s-value, which describes the entirety of the interacting population and whose s-value lies in between the true s-values of monomer and dimer. The value of this weighted average s-value changes with concentration, e.g. at a higher concentration more dimer would be expected, so the observed peak in the $c(s)$ distribution would shift to a greater s-value, corresponding to a higher proportion of dimer. If the exchange is intermediate, distinct peaks can be observed for each species present, but their s-values may also shift with changing concentration.

A lower temperature slows down the chemical exchange between monomer and other oligomeric species and the reaction becomes kinetically limiting, allowing two separate sedimenting boundaries to form due to lack of local equilibria. Therefore, separate species, can be observed in the $c(s)$ distribution. The species may also have weighted-average s-values which vary with concentration. High concentrations, as well as low diffusion coefficients, e.g. if proteins are elongated, can give rise to sharp sedimenting boundaries (Schachman, 1959 & Cann *et al.*, 1970). This can allow individual peaks for each species to be observed despite the exchange occurring on the timescale of the experiment (Mavaddat *et al.*, 2000).

3.4.2 Sedimentation velocity vs. sedimentation equilibrium

The two AUC methods provide different yet complementary information on a solution of sedimenting macromolecules (Cole et *al.*, 2008). SV is a hydrodynamic method, providing information on the number of species present, their molecular weight, their shape and also kinetic information on interactions between them. SE is a thermodynamic method which can provide accurate information on the number of species present, their molecular weight

independent of their shape, their stoichiometry and also information on solution non-ideality.

While it may be presumed that reversible interactions which occur on the timescale of the sedimentation cannot be studied by SV, this method may be the only possibility if the sample is unstable or fails to reach equilibrium in a SE experiment (Cole *et al.*, 2008). Furthermore, SE lacks the excellent resolution of SV and requires a stricter sample homogeneity and purity. A greater number of species also makes it progressively difficult to fit SE concentration gradients. Finally, SE experiments last much longer than SV experiments and are thus disfavoured if the protein is unstable.

3.4.3 Absorbance vs. interference optics

The choice of detecting system largely depends on the species under study (reviewed in Balbo & Schuck, 2005). Absorbance optics has the advantage of being selective to proteins (in the presence of non-absorbing components) and it is also not as sensitive to buffer mismatch between the reference and sample cells. In addition, lower protein concentrations are required than when using interference optics. However, common problems arise if a particular buffer component absorbs very strongly or if the protein under study has no aromatic residues and therefore an extinction coefficient of zero at 280 nm. An additional problem is that a very high concentration can lead to deviations from the linearity between sample concentration and absorbance.

In contrast, interference optics exhibit unlimited linearity and can be used with proteins which lack aromatic amino acids, such as some of the HR1 domains. Strongly absorbing buffer components are also not important. Common problems include the need to use higher protein concentrations and also the need for very good buffer and meniscus matching to prevent noise in the data from the protein-buffer data subtraction. The recording time for each scan is however much faster than when absorbance optics are used.

3.4.4 Experimental conditions used

In this study, a rotor speed of 40000 rpm (An60Ti rotor) was chosen as the appropriate SV speed for proteins with size of 10-30 kDa which may form oligomers of much larger size. At higher speeds, e.g. 50000 or 60000 rpm, which are routinely used, there is the risk that high molecular weight species sediment very quickly and are poorly resolved when the data are fit. The temperature at which the experiments were conducted was chosen to be 4 ^{o}C, in an attempt to slow down the exchange between the various species in solution, allowing the accurate determination of the various oligomeric species present. In addition, two of the proteins are only stable at 4 ^{o}C. The experiments performed, the species observed, their corresponding *s*-values and molecular weights, and the fitted frictional ratios observed are outlined in table A3.1 (see Appendix 3).

The program SEDNTERP was used to calculate the buffer density, buffer viscosity and protein partial specific volume based on the solvent and protein amino acid compositions.

3.4.5 PRK1 HR1a

For PRK1 HR1a, a sedimentation velocity experiment was performed at three concentrations (doubling dilution); 100 µM, 200 µM and 400 µM (Figure 3.20). The data were fit with a continuous *c(s)* distribution, and two species were observed. At 100 µM, 59.6% monomer (14.9 kDa) and 40.4% dimer (34.6 kDa) were observed, with respective *s*-values of 0.974 and 1.64. At 200 µM, 41.4% monomer (15.6 kDa) and 58.6% dimer (34.4 kDa) were observed, with *s*-values 1.04 and 1.77. At 400 µM, 26.5% monomer (19.4 kDa) and 73.5% dimer (37.5 kDa) were observed, with *s*-values 1.15 and 1.79. These results indicate that dimerisation is concentration-dependent and that the K_d of dimerisation is approximately 150 µM. The masses had to be scaled according to monomer-dimer, i.e. assume the smaller species is the monomer and identify the larger species as the dimer based on the 1:2 ratio of the MW estimated when fitting the data.

The fitted frictional ratio, $\frac{f}{f_0}$, was overestimated (leading to overestimated calculation of MW) and it is predicted, using the program HYDROPRO (Ortega *et al.*, 2011), that it should be less than the average value of 1.3 found here.

The *s*-value of the monomer is very similar to the *s*-value of the HR1b monomer (see section 3.4.6), and the mass ratio of the two species is very close to 2, so it seems reasonable that the two species are monomer and dimer. There is also monomer-dimer exchange which is indicated by the shifting *s*-values of the monomer peak, suggesting that dimerisation is reversible and that monomer and dimer exist in a dynamic equilibrium.

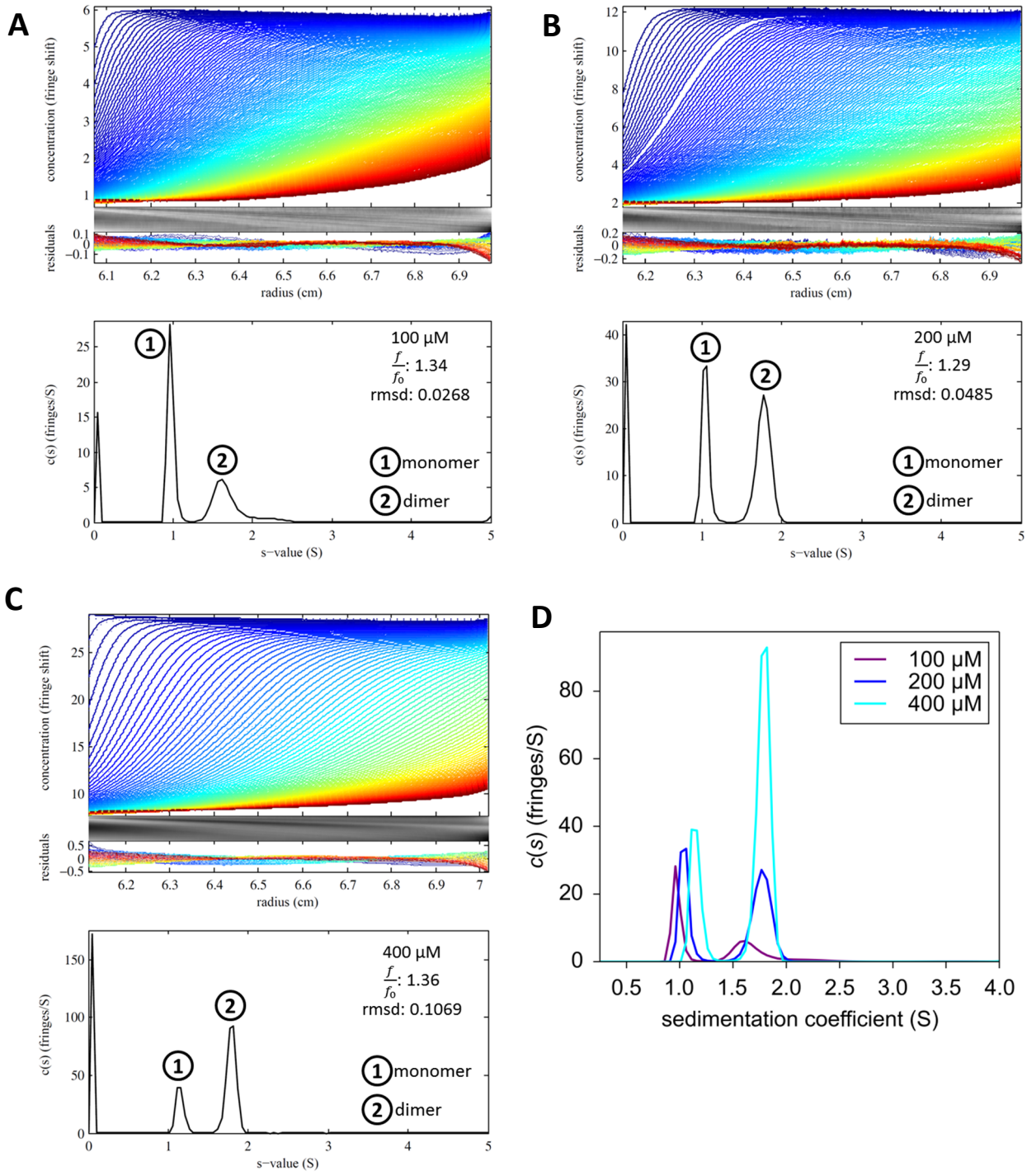


Figure 3.20 PRK1 HR1a sedimentation velocity experiments. (**A**)-(**C**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK1 HR1a at (**A**) 100 µM (**B**) 200 µM (**C**) 400 µM (**D**) Overlay of the c(s) distributions.

3.4.6 PRK1 HR1b

PRK1 HR1b sedimentation velocity experiments were performed at three concentrations of a doubling dilution, 200 µM, 400 µM and 800 µM (Figure 3.21). The data were fit with a continuous *c(s)* distribution, and a major species was observed at an average *s*-value of 0.73. The fitted frictional ratio is 1.42, which agrees with the known structure of HR1b i.e. that of an elongated coiled coil. The molecular weight from the fitted data is 10.3 kDa, which is very close to the actual molecular weight of 9.1 kDa. HR1b is monomeric as expected, agreeing with the NMR data recorded in our lab on this protein. A species at an *s*-value of 2 was also observed but this reflected only 3% of the population of molecules. This is not an oligomer as its proportion does not increase when the concentration is doubled.

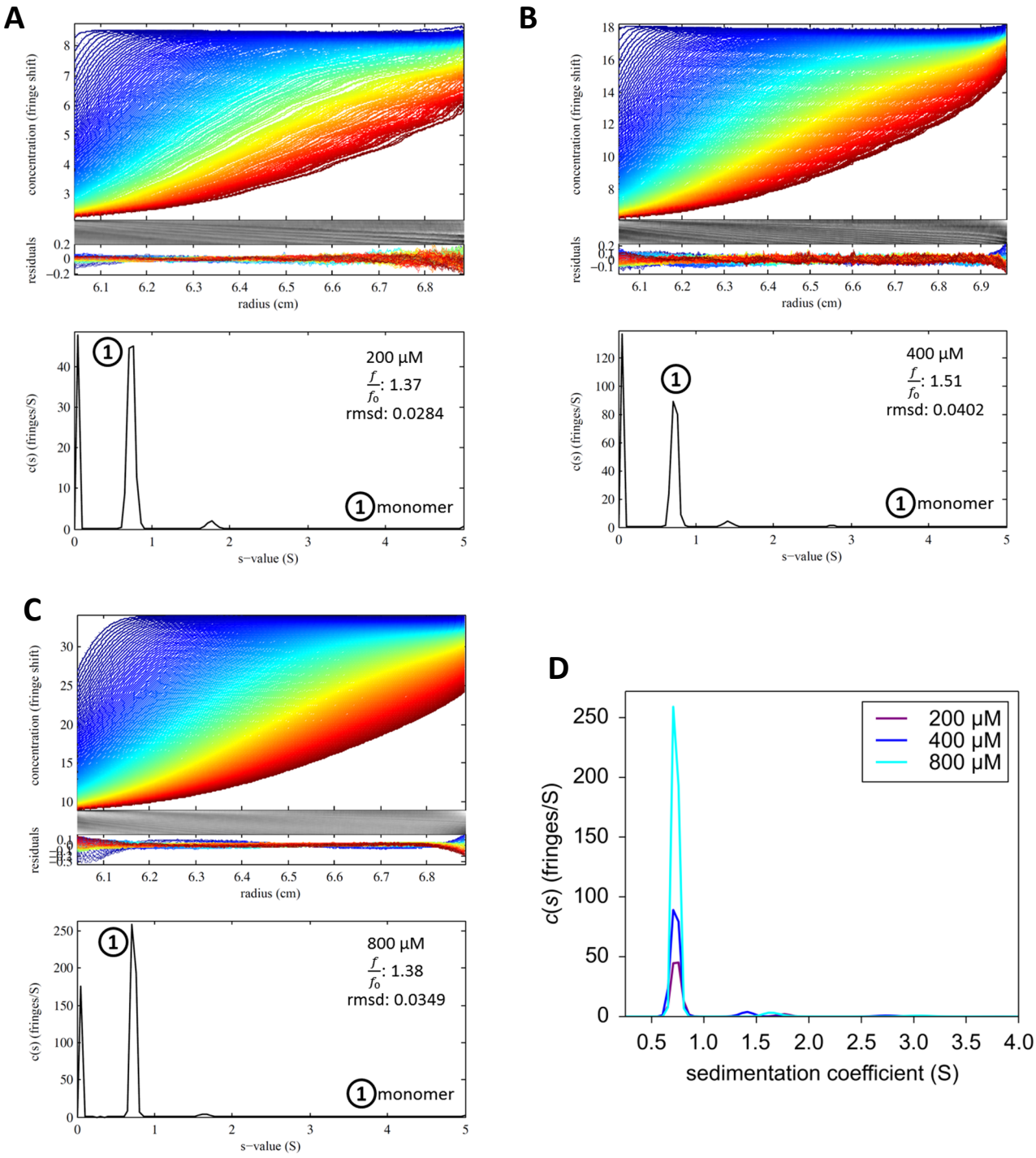


Figure 3.21 PRK1 HR1b sedimentation velocity experiments. (**A**)-(**C**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK1 HR1b at (**A**) 200 µM (**B**) 400 µM (**C**) 800 µM (**D**) Overlay of the c(s) distributions.

3.4.7 PRK1 HR1c

A sedimentation velocity experiment was performed for PRK1 HR1c at two concentrations, 400 μM and 800 μM (Figure 3.22). The data were fit with a continuous *c(s)* distribution, and a single species, the monomer (13.2 kDa) was observed, with an average *s*-value of 0.77, which is very similar to that of PRK1 HR1b. No higher order oligomers were observed at the concentrations chosen. The fitted frictional ratio, an average of 1.59 from the two experiments, indicates an elongated protein, which is consistent with HR1c being expected to be structured as an elongated coiled coil. However this frictional ratio also suggests that HR1c is slightly more elongated than HR1a and HR1b. This is unlikely to be due to differences in rigidity, as HR1b and HR1c have similar melting temperatures, but may be due to the unique C-terminal helix of HR1c.

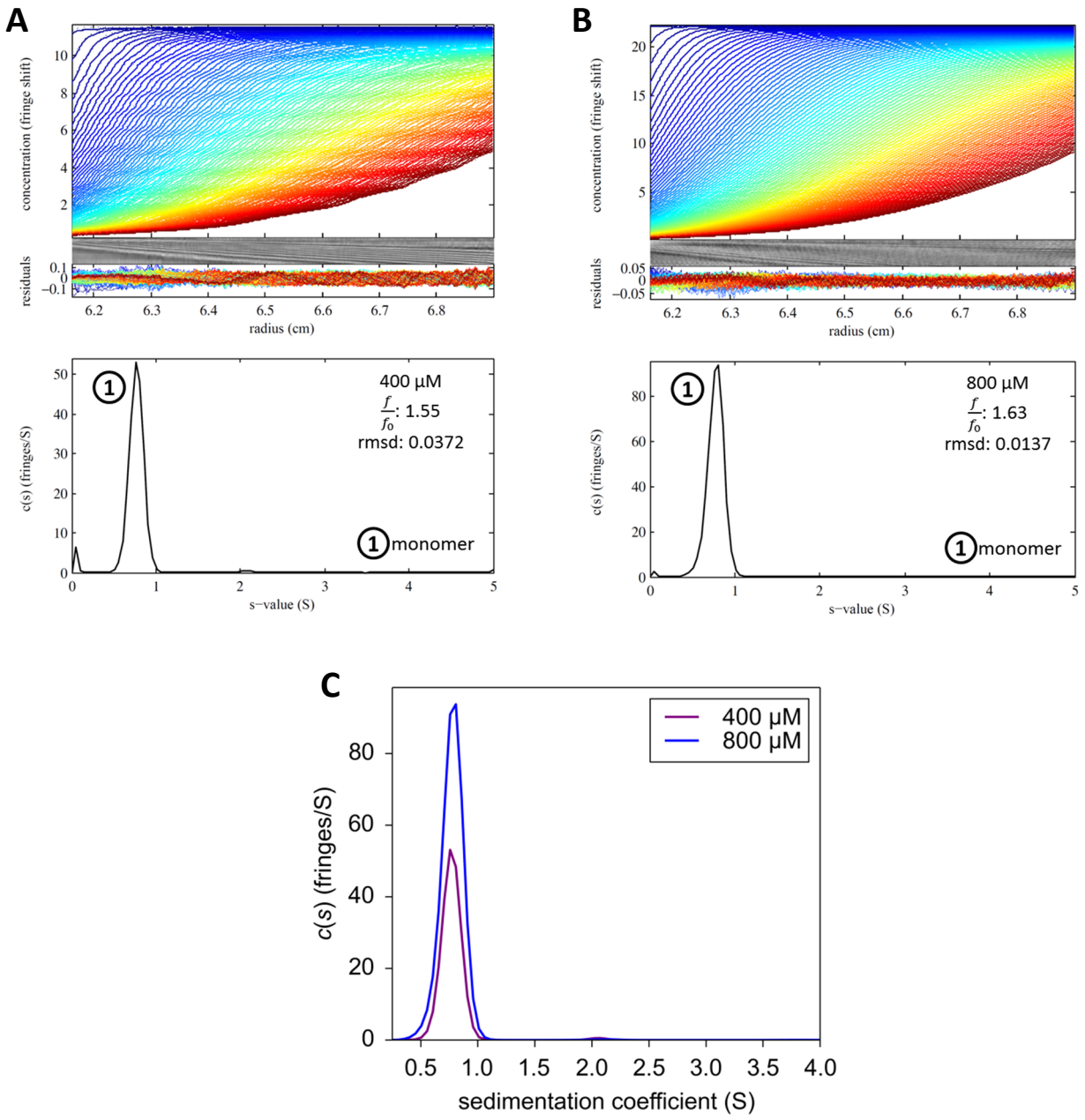


Figure 3.22 PRK1 HR1c sedimentation velocity experiments. (**A**)-(**B**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK1 HR1c at (**A**) 400 µM (**B**) 800 µM (**C**) Overlay of the c(s) distributions.

3.4.8 PRK1 HR1ab

For PRK1 HR1ab, a sedimentation velocity experiment was performed at three concentrations (doubling dilution); 100 µM, 200 µM and 400 µM (Figure 3.23). The data were fit with a continuous *c(s)* distribution, and two species were observed. At 100 µM, 56.6% monomer (18.7 kDa) and 43.4% dimer (44.7 kDa) were observed. At 200 µM, 39.1% monomer (22.5 kDa) and 60.9% dimer (51.0 kDa) were observed. At 400 µM, 22.5% monomer (28.5 kDa) and 77.5% dimer (61.9 kDa) were observed. At this highest concentration the frictional ratio is slightly overestimated as a result of solution non-ideality, which in turn leads to the calculated molecular weights being overestimated too. Collectively, these results indicate that dimerisation is concentration-dependent and that the K_d of dimerisation is approximately 150 µM. Monomer-dimer exchange is indicated by the shifting *s*-values of the dimer peak. However, the *s*-value doesn't shift as much as in HR1a, suggesting that the monomer-dimer exchange happens on a slower time-scale. Both the K_d of dimerisation, as well as the monomer-dimer exchange, are consistent with the PRK1 HR1a data, suggesting that HR1a may have a role as an oligomerisation domain in PRK1. No higher order oligomers were observed at the concentrations chosen, suggesting that the oligomerisation stops at dimer, as in PRK1 HR1a. Additionally, the frictional ratio is lower than that of the single HR1b domain, suggesting that the di-domain is more globular in comparison.

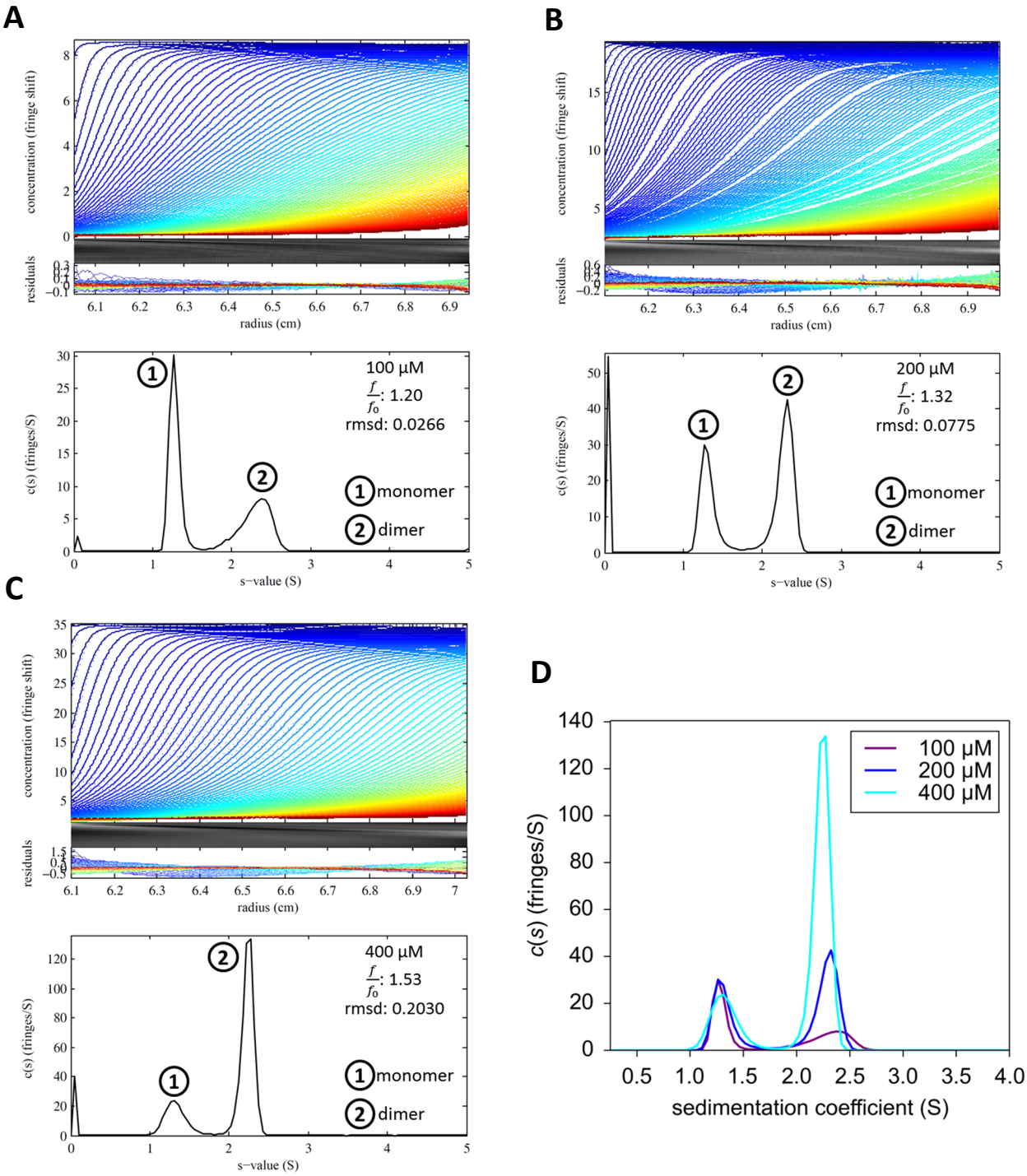
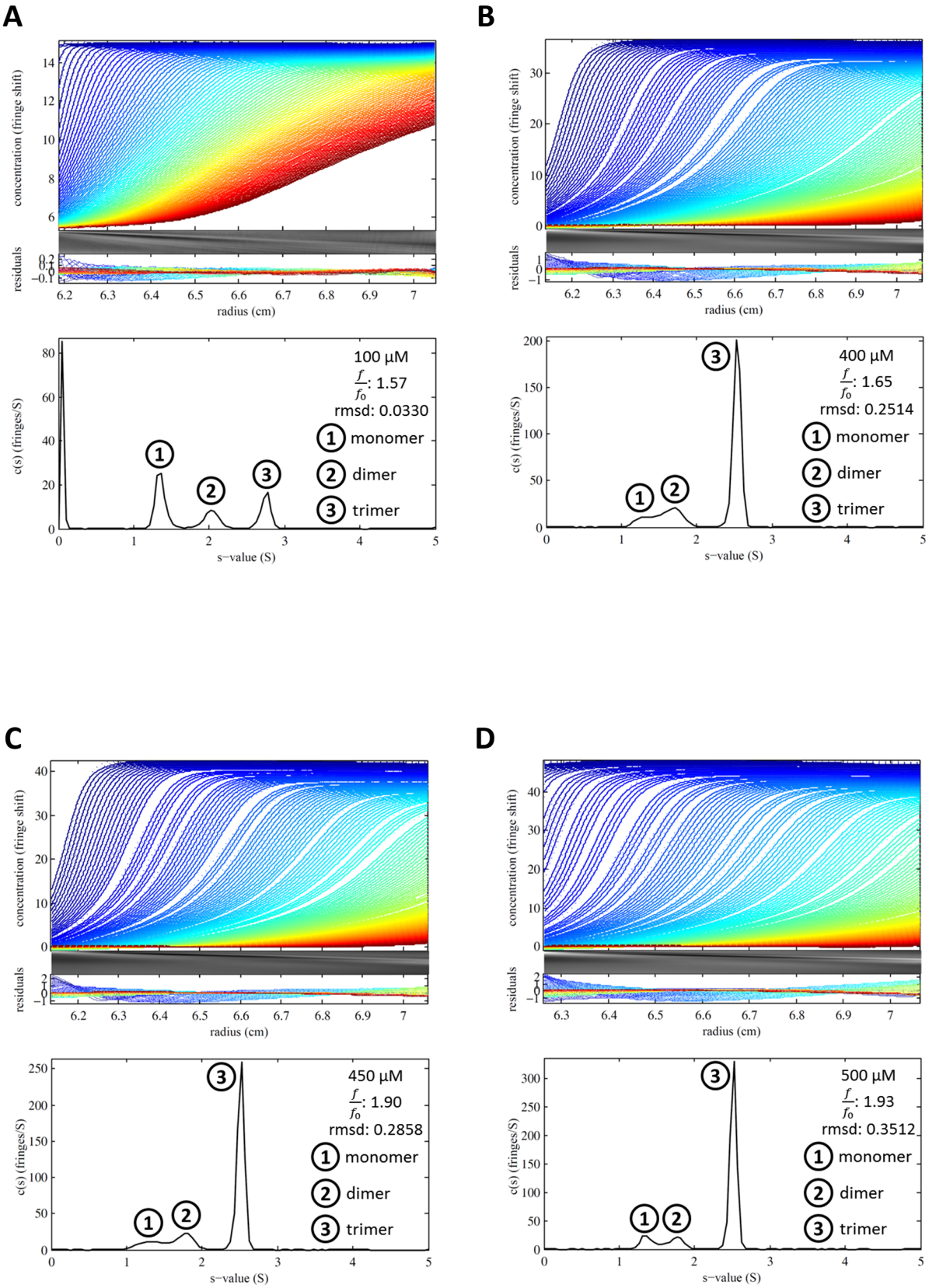


Figure 3.23 PRK1 HR1ab sedimentation velocity experiments. (**A**)-(**C**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK1 HR1ab at (**A**) 100 µM (**B**) 200 µM (**C**) 400 µM (**D**) Overlay of the c(s) distributions.

3.4.9 PRK1 HR1abc

To study the PRK1 HR1abc tri-domain, sedimentation velocity experiments were set up at a range of concentrations (Figure 3.24). The data were fit with a *c(s)* distribution, and three species were observed. At 100 µM, these species have *s*-values of 1.36, 2.02 and 2.74. These were best fit to a monomer-dimer-trimer system, which masses of 30.6 kDa, 55.5 kDa and 87.6 kDa, respectively. The abundances were 47.8% monomer, 23% dimer and 29% trimer. The molecular weight of PRK1 HR1abc is 33.1 kDa, so this indicates that the fitted global frictional ratio, 1.57, matches the true frictional ratio of each of the species, which do not become more globular upon oligomerisation. The dimer, and to a lesser extent the trimer, have slightly lower molecular weight than expected as a result of boundary broadening caused by dynamic exchange between the species. At 400 µM, there was 9.1% monomer (31.0 kDa), 18.2% dimer (45.8 kDa) and 72.7% trimer (83.6 kDa). The respective *s*-values were 1.30, 1.69 and 2.53. At 450 µM, there was 10.8% monomer (38.7 kDa), 15.5% dimer (59.9 kDa) and 73.7% trimer (102 kDa). The respective *s*-values were 1.30, 1.77 and 2.51. At 500 µM, there was 10.7% monomer (43.2 kDa), 10.8% dimer (60.9 kDa) and 78.5% trimer (104 kDa). The respective *s*-values were 1.38, 1.75 and 2.52. The shifting *s*-values of all 3 species from low to high concentrations are indicative of dynamic exchange between monomer-dimer-trimer. The increase in the trimer population at a higher concentration indicates that oligomerisation is concentration-dependent and the species exist with linked equilibria between them. At the concentrations chosen, no higher order oligomers were observed.


A
concentration (fringe shift)
residuals
radius (cm)
c(s) (fringes/S)
s−value (S)
100 µM
f/f0: 1.57
rmsd: 0.0330
1 monomer
2 dimer
3 trimer
B
concentration (fringe shift)
residuals
radius (cm)
c(s) (fringes/S)
s−value (S)
400 µM
f/f0: 1.65
rmsd: 0.2514
1 monomer
2 dimer
3 trimer
C
concentration (fringe shift)
residuals
radius (cm)
c(s) (fringes/S)
s−value (S)
450 µM
f/f0: 1.90
rmsd: 0.2858
1 monomer
2 dimer
3 trimer
D
concentration (fringe shift)
residuals
radius (cm)
c(s) (fringes/S)
s−value (S)
500 µM
f/f0: 1.93
rmsd: 0.3512
1 monomer
2 dimer
3 trimer

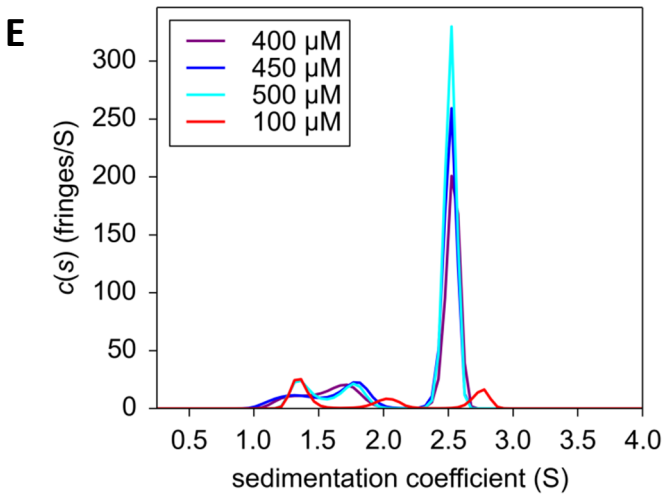


Figure 3.24 PRK1 HR1abc sedimentation velocity experiments. (**A**)-(**D**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK1 HR1a at (**A**) 100 µM (**B**) 400 µM (**C**) 450 µM (**D**) 500 µM (**E**) Overlay of the c(s) distributions.

3.4.10 PRK3 HR1a

Sedimentation velocity experiments were set up at three different concentrations (Figure 3.25). The data were fit with a *c(s)* distribution and two species were observed, which best fit to a monomer-dimer model. At 100 µM, 89.9% monomer (12.0 kDa) and 10.1% dimer (31.8 kDa) were observed, with *s*-values 0.782 and 1.49. At 200 µM, 78.1% monomer (11.7 kDa) and 21.9% dimer (24.0 kDa) were observed, with *s*-values 0.713 and 1.15, while at 400 µM, 71.9% monomer (11.3 kDa) and 28.1% dimer (22.3 kDa) were observed, with *s*-values 0.729 and 1.15. No higher molecular weight species were detected. The K_d of dimerisation is approximated to 1 mM, which is more than 6 times higher than the K_d of dimerisation of PRK1 HR1a. The fitted frictional ratio indicates an elongated protein, as predicted, and is also larger than the frictional ratio of PRK1 HR1a, possibly due to PRK1 HR1a being slightly more globular due to the presence of its N-terminal helix, which is absent from PRK3 HR1a. The slightly shifted *s*-values at 100 µM are a result of a high signal-to-noise ratio at this low concentration.

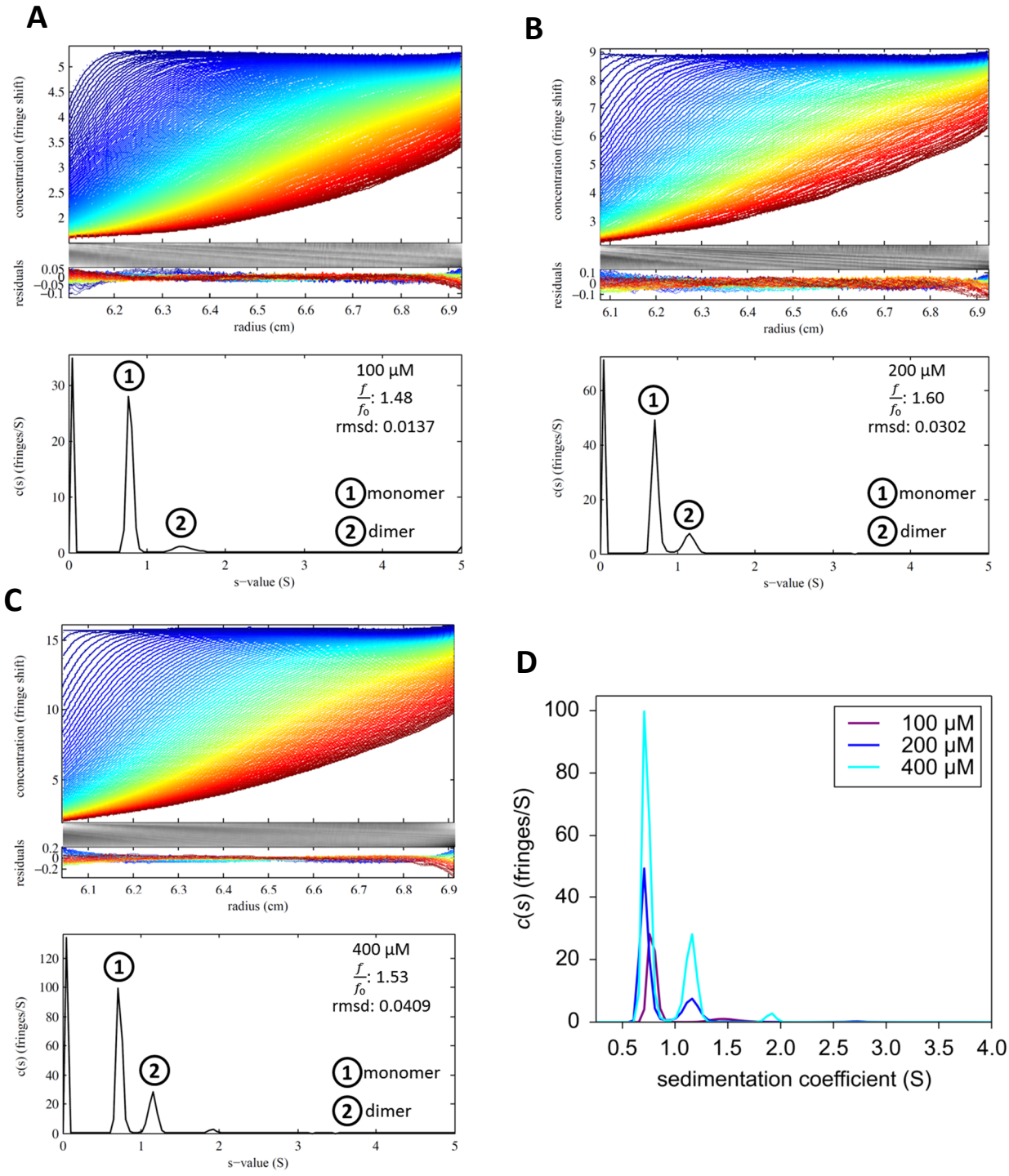


Figure 3.25 PRK3 HR1a sedimentation velocity experiments. (**A**)-(**C**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK3 HR1a at (**A**) 100 µM (**B**) 200 µM (**C**) 400 µM (**D**) Overlay of the c(s) distributions.

3.4.11 PRK3 HR1b

PRK3 HR1b oligomerisation was also investigated with a sedimentation velocity experiment (Figure 3.26). The data were fit with a *c(s)* distribution and two species were observed, which best fit to a monomer-dimer system. At 200 µM, 90.3% monomer (11.8 kDa) and 9.7% dimer (25.2 kDa) were observed, with corresponding *s*-values of 0.716 and 1.19. At 400 µM, 85.1% monomer (12.7 kDa) and 14.9% (28.8 kDa) dimer were observed, with *s*-values of 0.727 and 1.26, respectively. No further oligomerisation was detected. The K_d of dimerisation is approximated to 2 mM, which is the highest dimerisation K_d of any of the HR1 domains studied here.

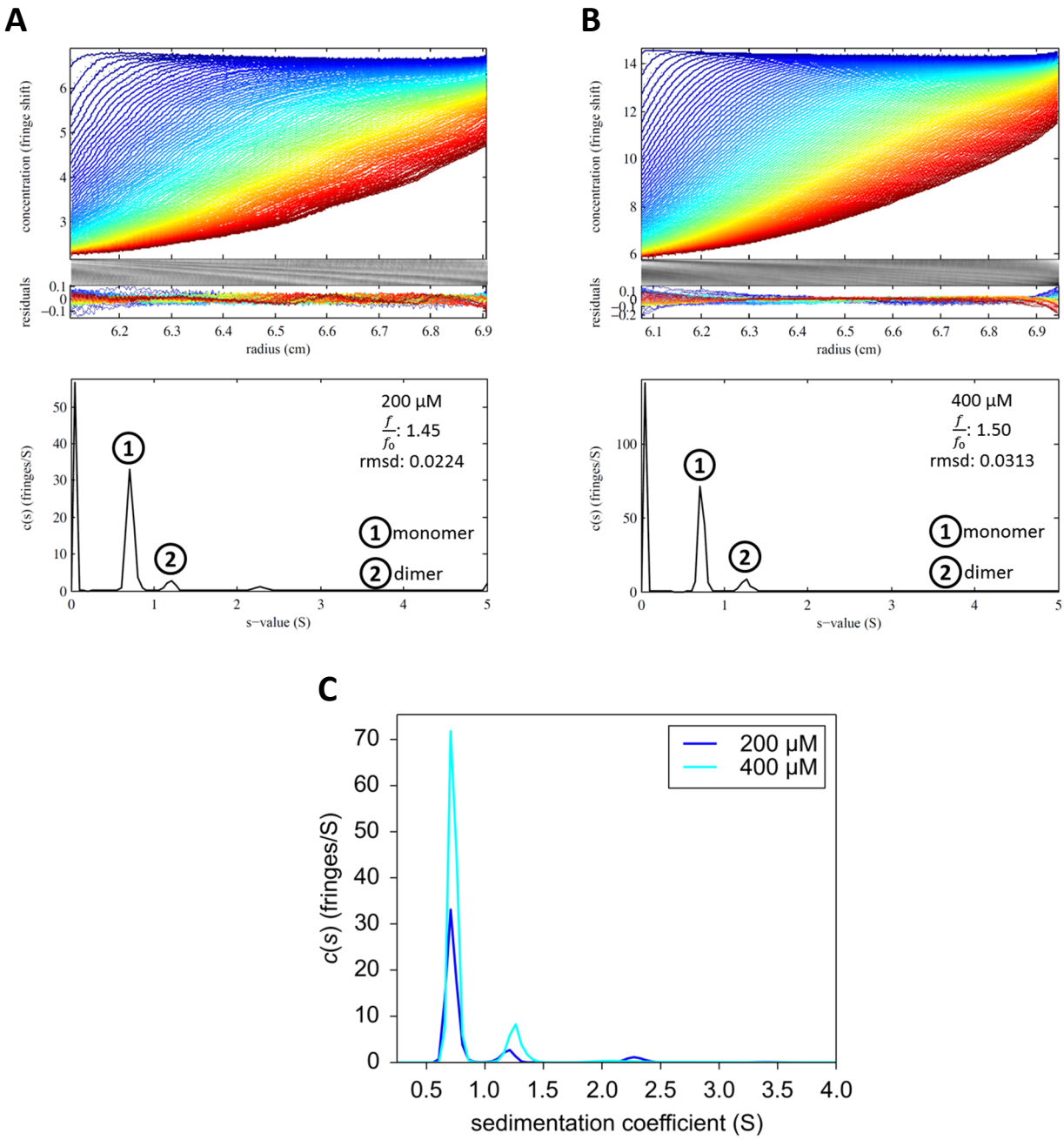


Figure 3.26 PRK3 HR1b sedimentation velocity experiments. (**A**)-(**B**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK3 HR1b at (**A**) 200 µM (**B**) 400 µM (**C**) Overlay of the c(s) distributions.

3.4.12 PRK3 HR1c

The self-association of PRK3 HR1c was analysed by sedimentation velocity experiments conducted at a range of concentrations with a doubling dilution (Figure 3.27). The data were fit with a *c(s)* distribution, and three species were observed, which were best fit to a monomer-dimer-tetramer model. At 100 µM, there was 49.1% monomer (12.7 kDa), 44.7% dimer (20.5 kDa) and 6.2% tetramer (48.6 kDa). The respective *s*-values were 0.666, 0.949 and 1.68. At 200 µM, there was 36.4% monomer (10.9 kDa), 56.8% dimer (23.2 kDa) and 6.8% tetramer (46.9 kDa). The respective *s*-values were 0.606, 1.02 and 1.63. At 400 µM, there was 34.7% monomer (14.2 kDa), 49.9% dimer (25.9 kDa) and 15.4% tetramer (41.4 kDa). The respective *s*-values were 0.731, 1.10 and 1.51. Therefore the data suggests that PRK3 HR1c exists in a linked equilibrium of monomer-dimer-tetramer. This is in contrast to PRK1 HR1c which is monomeric. The shifting *s*-values of all 3 species are indicative of dynamic exchange between monomer-dimer-tetramer. The K_d between monomer and dimer is estimated to be ~150 µM, which is very similar to the dimerisation K_d of PRK1 HR1a and PRK1 HR1ab. The increase in the tetramer population at a higher concentration also indicates that oligomerisation is concentration-dependent. At the concentrations chosen, no higher order oligomers were observed.

The frictional ratio is the highest than any of the single HR1 domains studied here and slightly higher than the frictional ratio of PRK1 HR1c. This could account for one HR1c being monomeric and the other being oligomeric, however PRK1 HR1a, which has the lowest frictional ratio, dimerises with similar affinity. Therefore the frictional ratio cannot be used to predict whether an HR1 domain oligomerises or not.

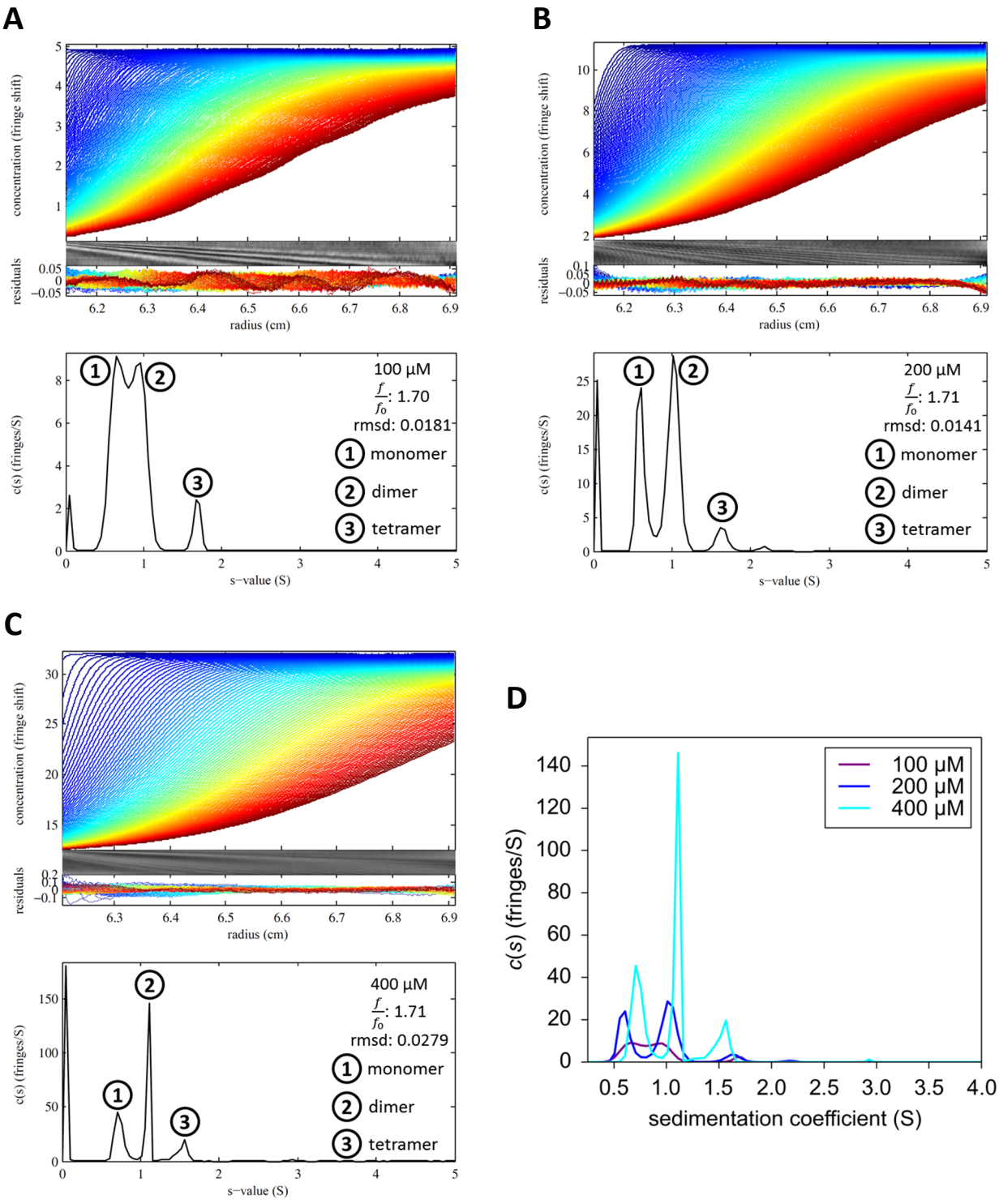


Figure 3.27 PRK3 HR1c sedimentation velocity experiments. (**A**)-(**C**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK3 HR1c at (**A**) 100 µM (**B**) 200 µM (**C**) 400 µM (**D**) Overlay of the c(s) distributions.

3.4.13 PRK3 HR1ab

A sedimentation velocity experiment was performed with the PRK3 HR1ab di-domain (Figure 3.28). The data were fit with a *c(s)* distribution and two species were observed, which best represent a monomer-dimer system. At 50 µM PRK3 HR1ab, 91.4% monomer and 8.6% dimer were observed, with *s*-values 1.08 and 1.55. No further oligomers were observed. The K_d of dimerisation is estimated to be approximately 500 µM, indicating a twice as strong dimerisation as the single PRK3 HR1a domain and a four times as strong dimerisation as in the case of PRK3 HR1b, suggesting that in the di-domain the dimerisation may involve both of these single domain oligomerisation interfaces. However, in comparison to PRK1 HR1ab, the K_d is still more than three times higher.

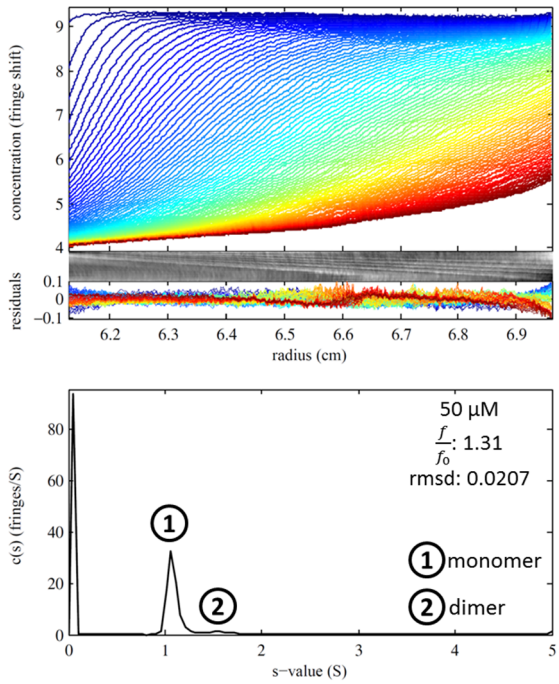


Figure 3.28 PRK3 HR1ab sedimentation velocity experiment. Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profile of PRK3 HR1ab at 50 µM.

3.4.14 PRK3 HR1abc

Sedimentation velocity experiments were set up at a range of PRK3 HR1abc concentrations and the sedimentation was monitored with interference (Figure 3.29). The data were fit with a *c(s)* distribution, and four species were observed. At 50 µM, these species have *s*-values of 1.42, 2.10, 2.78 and 4.04. These were best fit to a monomer-dimer-trimer-larger oligomer system, with masses of 20.8 kDa, 37.3 kDa, 58.1 kDa and 98.7 kDa, respectively. The molecular weight of PRK3 HR1abc is 33 kDa, so this indicates that the fitted frictional ratio, 1.07, is an underestimate assuming that the smallest species at *s*-value 1.42 is the monomer. In practice, this is a logical conclusion as a frictional ratio of 1.07 would indicate species which are almost a perfect sphere. The fitted frictional ratio, $\frac{f}{f_0}$, was underestimated due to (a) the dominance of the buffer peak at *S*=0 (a result of slight buffer mismatch) and also (b) monomer, dimer, etc. exchange, which leads to a broadening of the boundary leading to an overestimate of the diffusion rate. Therefore the masses had to be scaled according to monomer, dimer, trimer and larger oligomer. The mass ratios of the three smallest species are very close to 1-2-3, so this seems reasonable. The identity of the larger oligomer is impossible to predict from this data (could be pentamer or hexamer). The monomer-dimer exchange is also indicated by the shifting *s*-values of the different species determined at higher concentrations. At 50 µM there was 13.5% monomer (20.8 kDa), 22.4% dimer (37.3 kDa), 55.8% trimer (58.1 kDa) and 8.3% larger oligomer (98.7 kDa). The respective *s*-values were 1.42, 2.10, 2.78 and 4.04. At 100 µM, there was 12.9% monomer (24.9 kDa), 29.1% dimer (48.8 kDa), 49.5% trimer (66.2 kDa) and 18.5% larger oligomer (113 kDa). The respective *s*-values were 1.51, 2.34, 2.89 and 4.13. At 200 µM, there was 5.1% monomer (21.1 kDa), 14.1% dimer (38.8 kDa), 54.5% trimer (60.4 kDa) and 26.3% larger oligomer (111 kDa), with respective *s*-values were 1.47, 2.23, 3.01 and 4.51. The shifting *s*-values of all four species are indicative of dynamic exchange between monomer-dimer-trimer-larger oligomer. The increase in the trimer and larger oligomer populations at higher concentrations indicates that oligomerisation is concentration-dependent and that the species exist under linked equilibria between them. At the concentrations chosen, no higher order oligomers were observed.

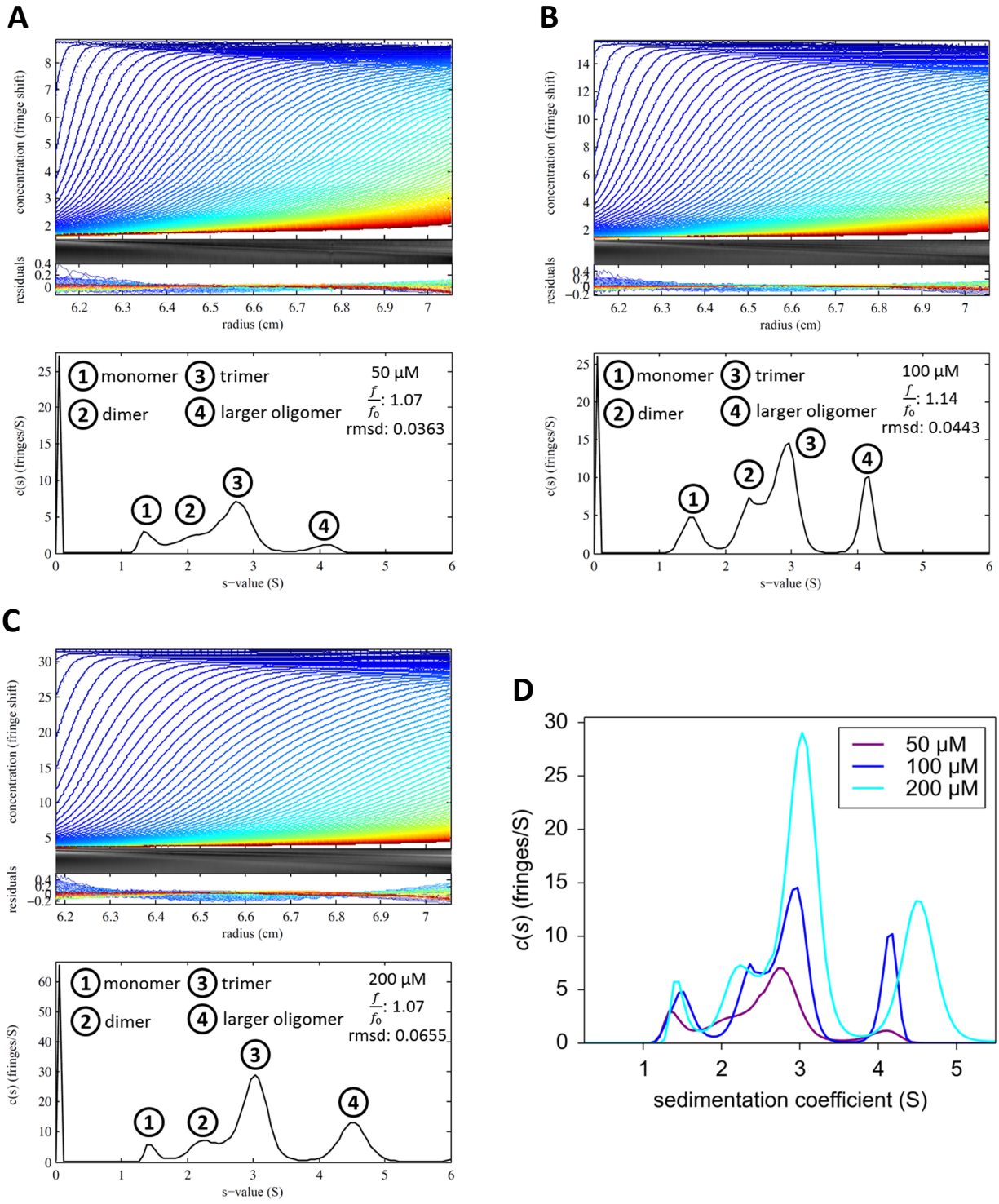


Figure 3.29 PRK3 HR1abc sedimentation velocity experiments. (**A**)-(**C**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of PRK3 HR1abc at (**A**) 50 µM (**B**) 100 µM (**C**) 200 µM (**D**) Overlay of the c(s) distributions.

3.5 Discussion and conclusions

The purification of PRK1 and PRK3 HR1 constructs was presented in this chapter. Gel filtration results led to the analysis of the self-association of these proteins by SV-AUC. PRK1 HR1a and PRK3 HR1c appear to be dimerisation domains for their respective PRKs. While in PRK1 the other two HR1 domains, HR1b and HR1c are monomeric and do not self-associate, PRK3 HR1a and PRK3 HR1b are weakly oligomerising, in addition to the HR1c domain. In addition to the inter-molecular interactions thought to be driven by these oligomerising HR1 domains, intra-molecular interactions in the context of the HR1abc tri-domains are also indicated by the more complex oligomerisation patterns. The oligomeric species of PRK1 and PRK3 HR1abc tri-domains detected in solution also suggest possible differential roles fulfilled by the different oligomeric states of the PRK N-terminal regulatory region. These roles may vary between PRK1 and PRK3, both by means of different oligomeric species formed but also possible structural differences between them. Nonetheless, this is the first time that HR1 domains are shown to be able to self-associate. In contrast, other HR1 domains, e.g. the HR1 domain from TOCA1, is monomeric (Watson, 2017; unpublished data).

All HR1 constructs purified and used in subsequent experiments are shown to be structured. This includes the two HR1c and two HR1abc domains from PRK1 and PRK3, which have never been purified from any construct previously in the lab. The thermal stability of the two HR1c domains was also tested, and while the T_m was similar for both proteins, the shape of the melting curve denotes structural differences between them.

3.5.1 Conclusions about PRK1 oligomerisation

As shown above, HR1b and HR1c are monomeric, while HR1a exists in an equilibrium of monomer-dimer. HR1ab also exists in a monomer-dimer equilibrium, with the K_d of dimerisation being similar to the HR1a dimerisation K_d, although the exchange between them may be slower, suggesting a possible interaction of HR1a with HR1b which is

independent of the HR1a-HR1a interaction. Therefore HR1a appears to act as an oligomerisation domain in PRK1. Unexpectedly, the HR1abc experiments suggest that the tri-domain exists in linked equilibria between monomer, dimer and trimer, which also appear to be dynamically exchanging between the different oligomeric states. This suggests both intra-molecular and inter-molecular interactions in the context of the tri-domain, mediated by HR1c, e.g. between HR1a and HR1c, or HR1b and HR1c and inter-molecular interactions, such as the HR1a-mediated dimerisation. The precise oligomerisation steps however, e.g. if monomer becomes dimer and then trimer or if monomer can form trimer directly, cannot be deduced from this data but are investigated further in chapter 5.

3.5.2 Conclusions about PRK3 oligomerisation

As shown above, HR1a, HR1b and HR1ab are all dimerising with very weak affinities, while HR1c exists in a dynamic equilibrium of monomer-dimer-tetramer and oligomerises more strongly. Therefore HR1c appears to act as the key oligomerisation domain in PRK3, while the other two HR1 domains may be important in stabilising the dimerisation in the full length protein. Furthermore, the HR1abc tri-domain exists in linked equilibria between monomer, dimer, trimer and a larger oligomer, which also appear to be dynamically exchanging between the different oligomeric states. This suggests both intra- and inter-molecular interactions in the context of the tri-domain, e.g. between HR1c and either HR1a or HR1b or both, and inter-molecular interactions, such as the HR1c-mediated dimerisation.

3.5.3 PRK1 and PRK3 oligomerisation comparison

4 species were observed in the case of PRK3 HR1abc while only 3 species were observed for PRK1 HR1abc. The largest oligomeric species is bigger than trimer, so it is exclusive to PRK3 HR1abc. Furthermore, the oligomerisation of PRK3 HR1abc involves higher affinity interactions, as indicated by the largest species already being the major species at low concentrations. While the PRK1 and PRK3 HR1abc tri-domains share oligomerisation

propensity, the particular number and nature of oligomeric species formed, the shape of these oligomers (as indicated by the fitted frictional ratios), as well as the strength of the interactions driving them (higher proportion of oligomeric species for PRK3 HR1abc when at the same concentration as PRK1 HR1abc), appear to be different. This may partly have to do with HR1a mediating the oligomerisation in PRK1 and HR1c mediating the oligomerisation in PRK3, although interestingly the dimerisation K_d of both PRK1 HR1a and PRK3 HR1c (the drivers of oligomerisation in their respective PRK) is similar.

The weak affinities recorded here may suggest that these interactions are not physiologically relevant, as these proteins will never be at such high concentrations in the cell to drive oligomerisation. However proteins often tend to be localised to a particular cell compartment, so their effective concentration is much higher. The PRKs do localise to the plasma membrane via their C2-like domain. It is therefore possible that recruitment to the membrane stimulates oligomerisation.

Furthermore, HR1a-mediated dimerisation would bring the other HR1 domains into proximity, therefore increasing their effective molarity even further and potentially driving HR1abc trimer formation.

A comparison of the stability of HR1 domains based on CD data from this chapter and Hutchinson *et al.* (2013) suggests that PRK1 HR1b and HR1c may be monomeric as a result of a higher melting temperature, in contrast to PRK1 HR1a which has a melting temperature that is lower by 6 ^{o}C and forms dimers (Figure 3.30). Rigidity does not seem to correlate with dimerisation when PRK1 HR1a and PRK3 HR1a are compared. Instead, PRK1 HR1a is the more rigid domain out of the two yet dimerises with 6 times higher the affinity. This may suggest that the unique N-terminal helix of PRK1 HR1a may be driving dimerisation. A look at the theoretical isoelectric point (pI) of each domain also does not establish a trend of dimerisation, in this case based on surface charge. Intra- and inter-molecular interactions in the PRKs could therefore be driven by weak, hydrophobic interactions.

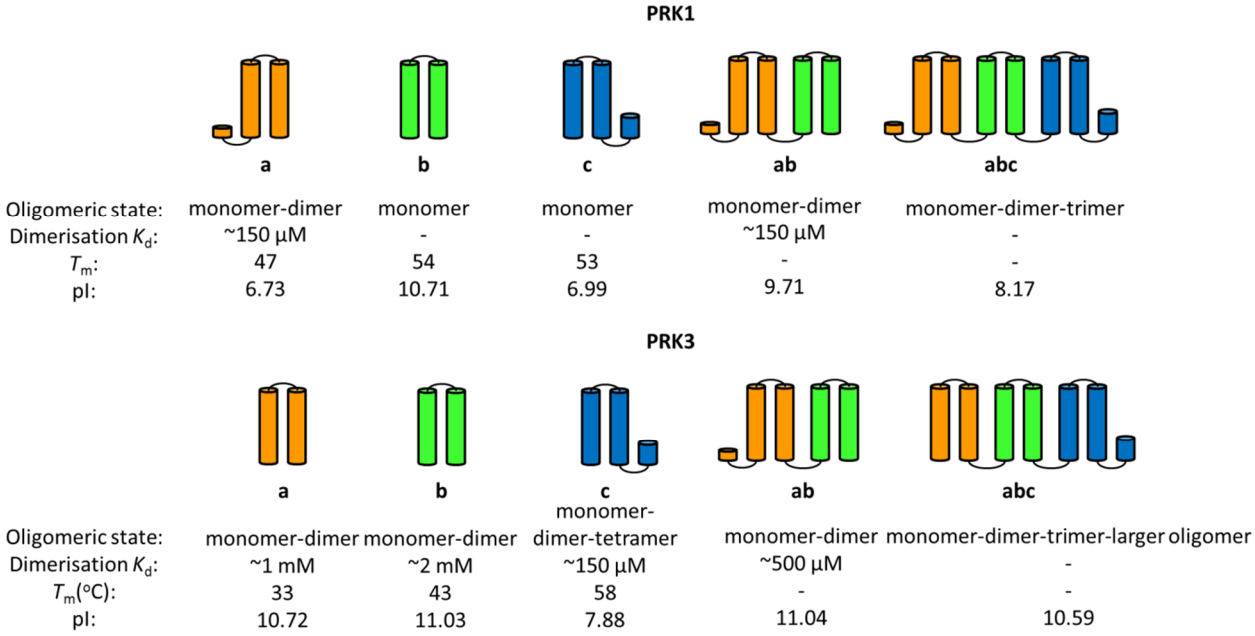


Figure 3.30 Comparison of biophysical properties of HR1 domains. The oligomeric state, dimerisation affinity, melting temperature and theoretical isoelectric point of each HR1 protein studied in this chapter is listed.

3.5.4 Importance of oligomerisation

This is the first time it has been shown that these HR1 domains are involved in such intra- and inter-molecular interactions and it is speculated that they are important in regulating PRK catalytic activity through dimerisation. The self-association of these domains may also serve to augment or abrogate the interactions of these domains with the Rho family of G proteins, which may in its own right affect the activation of the PRKs that are known to be activated directly by RhoA and Rac1 (Amano *et al.*, 1996; Lu & Settleman, 1999). Altered affinities for Rho G proteins may also affect the localisation of the PRKs and thus have an effect on the physiological role these proteins have in cells. Moreover, it has been shown previously that in the case of PRK2 there is oligomerisation mediated by a region C-terminal to the C2 domain (Bauer *et al.*, 2012), so it is also possible that there are multiple oligomerisation sites in the PRKs.

Weak, transient protein-protein interactions are widespread in cell signalling pathways and are key for their regulation, e.g. the SLAM glycoprotein dimerises with a K_d of 200 µM to 1 mM, which is thought to be important in transient cell-cell contacts made by leukocytes (Mavaddat *et al.*, 2000).

However, a lot remains unknown. Firstly, the structure of HR1c. Secondly, the structure of HR1abc which would show the interactions happening between the HR1 domains, both the intra-molecular and inter-molecular ones, and particularly the regions of the HR1 domains required for these interactions. Thirdly, the oligomerisation scheme, including K_d values and particularly association and dissociation rates (i.e. kinetic parameters), between HR1 domain interactions in the context of their different oligomeric states. Finally, the dimerisation interface of HR1a and the effect oligomerisation may have on GTPase binding. Some of these questions are dealt with in Chapters 4-6.

Chapter 4 Structural studies on the PRK1 HR1c domain

4.1 Introduction

PRK HR1 domains share ~30% sequence identity and ~60% sequence similarity. The structures of PRK1 HR1a and HR1b are known, and these proteins form antiparallel coiled coils. All other HR1 domains are also predicted to be coiled coils. Despite their similar structures, HR1 domains have different thermal stabilities, which may partly explain why they differentially interact with G proteins (Hutchinson *et al.*, 2013). Some of the HR1 domains are either known or predicted to have additional structural components. In particular, PRK1 HR1a has an N-terminal helix consisting of 5 amino acids folded in one helical turn whose role remains unknown (whether HR1a is truly structured in solution has not been proved yet; the only evidence is from a crystal structure in complex with RhoA (Maesaki *et al.*, 1999)). Similarly, HR1c domains from all three PRKs are predicted to have a C-terminal helix which is much shorter than the two main helices. Whether this C-terminal helix truly exists in solution also remains unknown. It has been shown that PRK1 HR1c does not interact with RhoA and no other GTPase interacting partners are known (Flynn *et al.*, 1998).

PRK1 HR1c has been successfully purified and initial structural studies by CD suggest that it is folded. The protein is also stable at room temperature and AUC shows that it is monomeric even at high concentrations. All of the above facilitate studies by Nuclear Magnetic Resonance (NMR) spectroscopy, which requires a monodisperse sample that is stable for the long duration of the experiments. The structure of the HR1c domain was investigated by NMR as described in this chapter. The structure will provide a means of understanding why the HR1c may exhibit differential stability, oligomerisation or binding to G proteins. In addition, the assigned NMR spectra of HR1c will provide a powerful tool for investigating possible interactions between the HR1 domains or with G proteins. An introduction to the basic NMR theory required to understand what the NMR signal is, and also how magnetisation is generated and transferred in the NMR experiments recorded, will be provided. This is to enable the basic understanding of what each of the different

experiments recorded shows. The peak assignment process in the different NMR spectra will also be briefly explained, as well as how this information enables structure calculations.

4.2 Nuclear Magnetic Resonance

4.2.1 The NMR signal

Nuclear Magnetic Resonance (NMR) is a widely used technique in the study of protein structure, with ~10% of the structures in the PDB being NMR solution structures. It is considered to be complementary to X-ray crystallography, and some proteins which fail to pack into ordered crystals have been found to generate well dispersed NMR spectra (Snyder *et al.*, 2005). NMR also has widespread uses in the study of protein-protein interactions, as the interacting interface can be determined even without the existence of a structure. Similarly, NMR is often used to study small molecules such as peptides or ligands, particularly those with therapeutic potential, and their interactions with macromolecules. NMR can provide a fingerprint for each spin-active nucleus and therefore provides excellent atomic resolution when studying macromolecules. A key advantage of NMR over other techniques is the ability to study protein dynamics and in particular protein conformations which are not detectable, e.g. very short-lived conformations which would unlikely be selected for in X-ray crystallography. While biological macromolecules are abundant in ^{1}H nuclei, the natural abundance of ^{15}N and ^{13}C is very low. At the µM and low mM concentrations used, the low abundance means that the signal from those nuclei will be very weak and lost in the noise. Consequently, biological macromolecules such as proteins have to be labelled with ^{15}N and ^{13}C, e.g. by growing the expression host in culture where the only sources of N and C are ^{15}N and ^{13}C, respectively.

When the NMR sample is placed in an NMR magnet, the magnetic moment of each nucleus interacts with the applied magnetic field, B_0. The energy of this interaction is dependent on the angle, θ, between the moment and the field direction. While there is an energetic advantage if the magnetic moments of all individual nuclei in our sample align with the

applied field, i.e. when θ=0, the thermal motion of the nuclei results in random magnetic moments which disrupts this alignment. However the disruption is not absolute, and a small excess of nuclei align with the applied magnetic field (Figures 4.1A-B). This overall net magnetisation of the entire NMR sample can be represented with a bulk magnetisation vector pointing along the z-axis on a set of axes and is referred to as the equilibrium magnetisation (Figure 4.1C).

If the equilibrium magnetisation is somehow tipped away from the z-axis (Figure 4.1D), there is precession about the field at frequency ω_0 in rads^{-1},

$$\omega_0 = -\gamma B_0 \qquad 4.1$$

where γ is the gyromagnetic ratio of the nucleus under study and B_0 is the strength of the magnetic field. This frequency is termed the Larmor frequency. In NMR experiments, it is this magnetisation vector precession that we detect. This is done by winding the detector coil round the x-axis. The precessing magnetisation will then 'cut' the coil as it cuts into the xy plane. The induced current results in free induction decay (FID), which is the NMR signal we record (Figure 4.1E). Fourier transformation of this signal produces the familiar signals or peaks in NMR spectra, at frequency ω_0.

NMR experiments require the magnetisation to move away from the z-axis and, for example, towards the x-axis. This is not possible as the magnetic field cannot be swiftly altered. The solution to this problem is to apply a small magnetic field, B_1, along the x-axis which oscillates at the Larmor frequency (Figure 4.2). When radiofrequency power (or pulse) is provided, an oscillating magnetic field along the x-axis is generated, which is enough to make the magnetisation precess in the yz plane. If we set the size of this oscillating field to $2B_1$, then we can depict two counter-rotating magnetic fields, each of size B_1, the x-components of which initially shrink towards zero as they move away from the x-axis but subsequently increase towards –x. This is impossible in the 'laboratory frame' as the magnetic field cannot be swiftly altered. Instead, the rotating frame is employed, where the coordinate system (the axes) are rotated at the same rate as B_1. Essentially, the B_1 then appears to be static and not oscillating, but the effect is the same as if it would have oscillated in the laboratory frame.

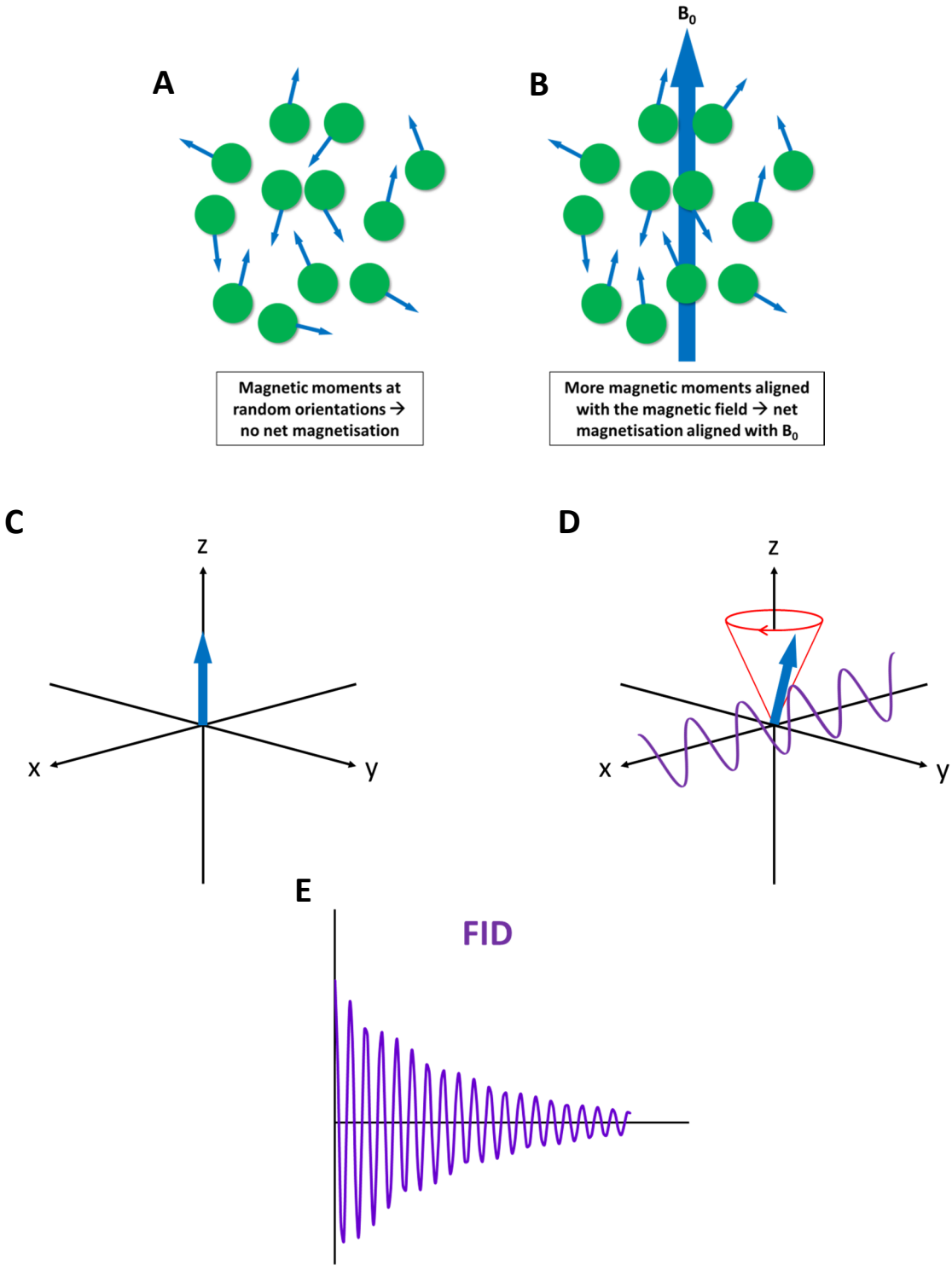


Figure 4.1 The NMR signal. (**A**) In an NMR sample, the magnetic moments of all nuclei are of random orientation. As a result, there is no net magnetisation. (**B**) The applied magnetic field, B_0, causes more moments to align with it. The net magnetisation is now aligned with the B_0 field. (**C**) The net magnetisation vector can be represented as being aligned on the z-axis of a xyz axis system. (**D**) When the magnetisation is tilted away from the z-axis, it precesses in the xy plane, about the z-axis. This means that it interacts at regular intervals with the receiver coil aligned on the x-axis. (**E**) This interaction gives the free induction decay (FID), which is the NMR signal.

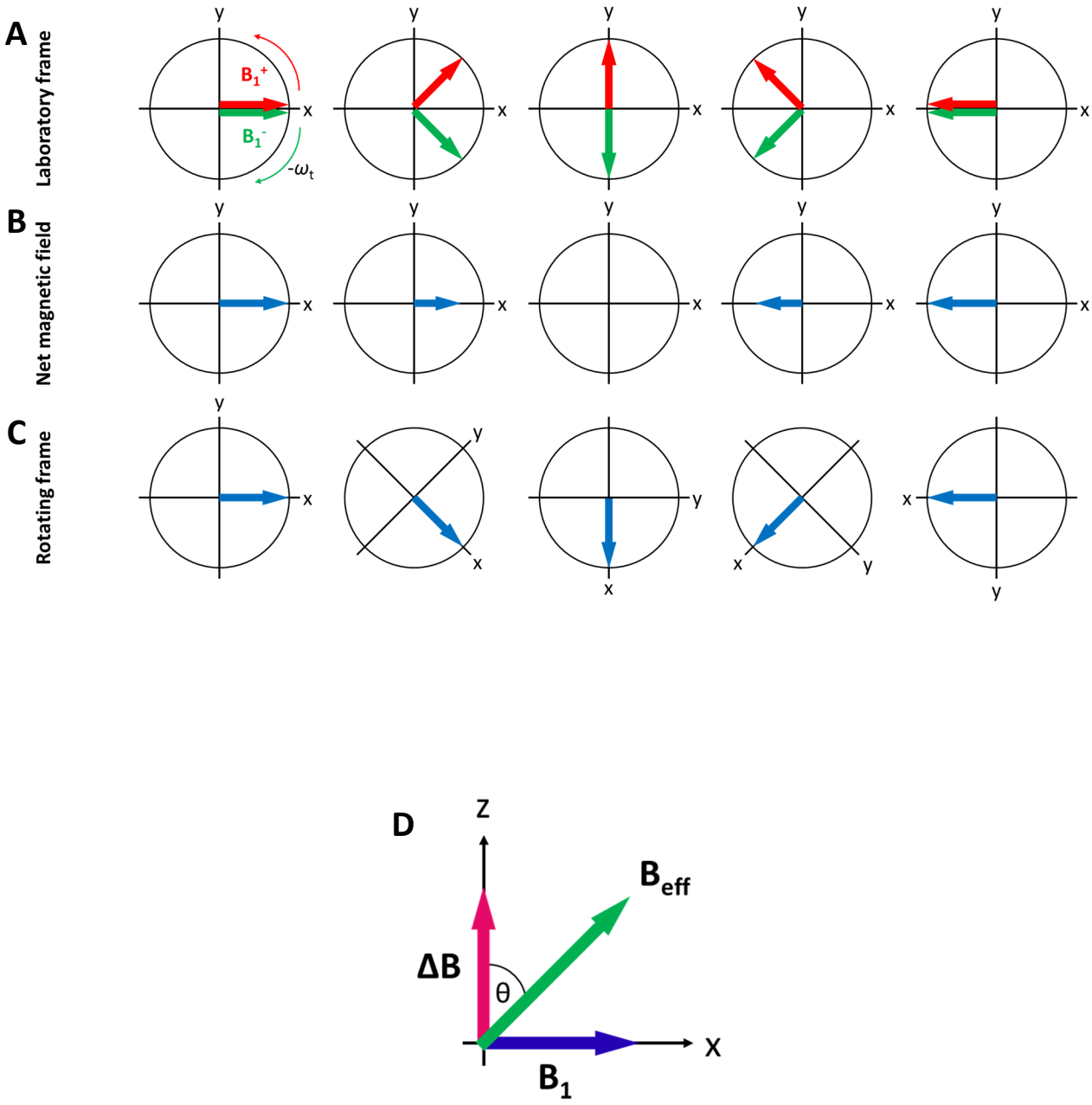


Figure 4.2 The rotating frame and the effective field. (**A**) Two counter-rotating fields, B^+_1 and B^-_1 in a fixed axis system known as the laboratory frame. (**B**) The sum of the B^+_1 and B^-_1 fields give a net magnetic field which oscillates about the x-axis. (**C**) In the rotating frame, the axis system is rotating at $-\omega_{tx}$ about the z-axis. The B^-_1 field appears to be static on the x-axis. (**D**) The effective field, B_{eff}, is the sum of the reduced field, ΔB, and the B1 field in the rotating frame.

If the rotating frame frequency is ω_{RF}, then the apparent Larmor frequency, also known as the offset (Ω), is the difference between the true Larmor frequency and the frequency of the rotating frame.

$$\Omega = \omega_0 - \omega_{RF} \qquad 4.2$$

Equation 4.1 describes how the Larmor frequency and the applied magnetic field are related. Substituting this equation and an analogous equation for the frequency of the rotating frame into equation 4.2 gives the relationship between the offset and the apparent magnetic field of the RF, ΔB.

$$\Omega = -\gamma \Delta B \qquad 4.3$$

The ΔB and B_1 magnetic fields add up to give the effective field, B_{eff}. It is this B_{eff} magnetisation vector which precesses and 'cuts' the detector coil to give the NMR signal. According to the above equation, if the Ω≈0, then ΔB is also ≈0, which means the B_{eff} aligns with the x-axis, i.e. the B_1 field has an effect.

NMR spectrometers can provide sufficient RF power to create pulses which cover the entire range of chemical shifts of a particular nucleus. In the basic pulse-acquire experiment (Figure 4.3), the equilibrium magnetisation is firstly allowed to align along the z-axis. The magnetisation then moves onto the –y axis following a 90° pulse on the x-axis. The magnetisation then precesses in the xy plane at frequency equal to the Ω. The chemical shift, δ, is often expressed in parts per million (ppm) rather than Hz,

$$\delta = \frac{\omega - \omega_{rv}}{\omega_{rv}} \times 10^6 \qquad 4.4$$

where ω_{rv} is the receiver or "operating" chemical shift of the spectrometer, e.g. 500 MHz.

A very commonly used pulse sequence is the spin echo. In this pulse sequence, a 180° pulse is added at the end of the basic pulse-acquire experiment, separating the sequence into two delays (2τ; τ=delay) and the signal is thus not acquired until after the second of the two delays. The effect that this has is that the magnetisation ends up along the same axis, independent of the length of the delay (τ) or the Ω. The magnetisation vector ends up where it started, which is why the 180° pulse is named a refocusing pulse, as chemical shift

evolution during the first delay is effectively negated by the second delay. We will see later that spin echoes are very useful when we want to interconvert in-phase and anti-phase magnetisation.

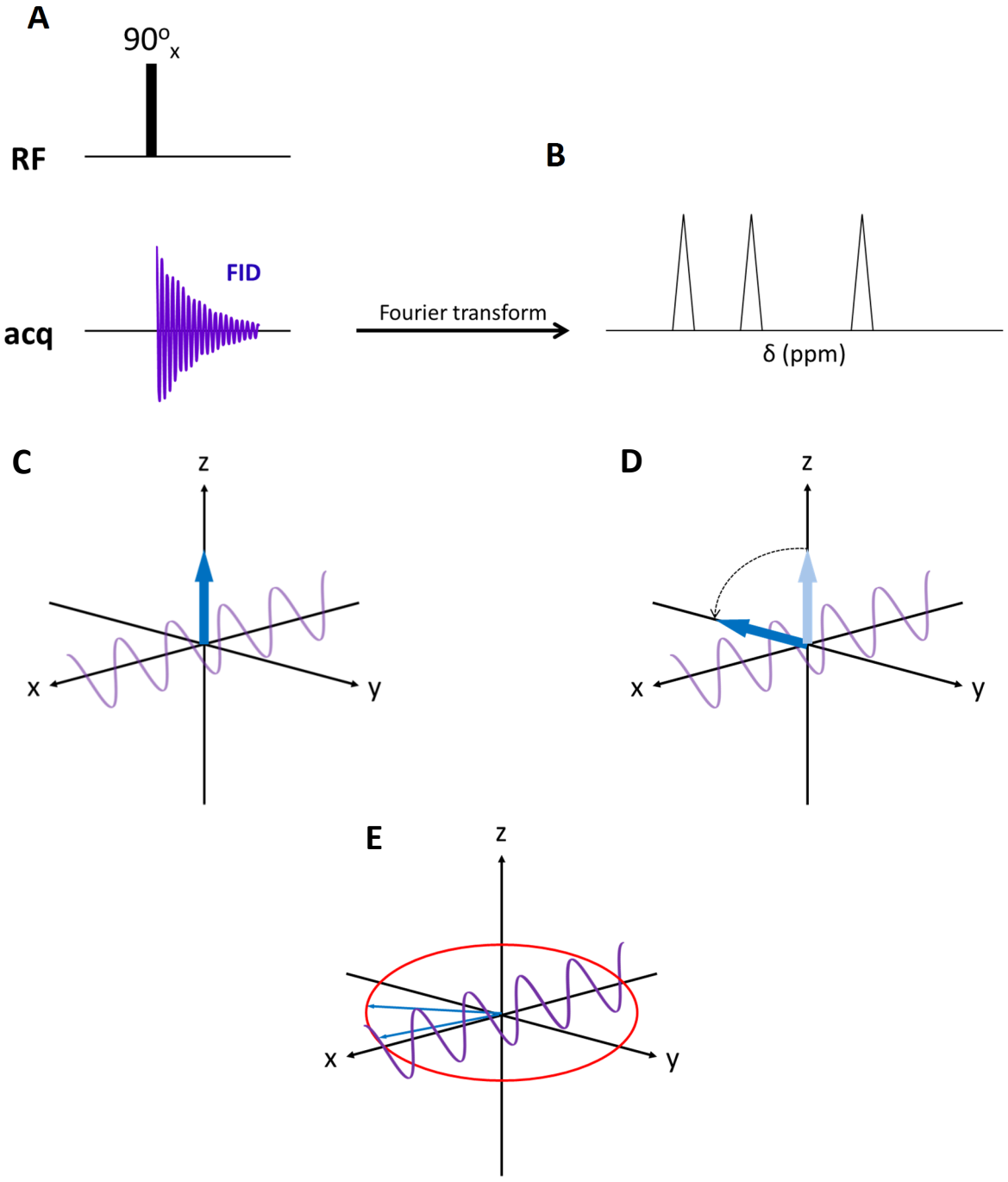


Figure 4.3 Pulse-acquire experiment. (**A**) The RF line contains pulses such as the 90° pulse on x shown here. The acquisition line shows the recorded signal. (**B**) The acquired FID is processed using a Fourier Transform to give a spectrum. This spectrum contains peaks which represent spin active nuclei and their position in the spectrum is at their Larmor frequency. This is typically expressed in chemical shift, δ, in units of ppm. (**C**) Before the 90° pulse, bulk magnetisation rests on the z-axis. (**D**) The 90° pulse on x rotates the magnetisation vector from the z-axis to the –y-axis. (**E**) The magnetisation precesses in the xy plane, cuts the receiver coil and the FID is acquired.

The above description of the basic NMR experiment only uses the vector model and a small number of equations. However as we move to more complicated experiments, involving two or more spins, we need to use the product operator approach to describe what is happening (Cavanagh *et al.*, 1998; Keeler, 2005). The elegance of this approach comes from the fact that by just following a few simple, quantum mechanical rules, even the most complex NMR experiments can be studied.

The energy of interaction between the spin and the magnetic field can be represented with the Hamiltonian operator

$$\widehat{H}_{one\ spin} = -\gamma B_0 \hat{I}_z \qquad 4.5$$

where $\widehat{H}$ is the Hamiltonian operator of the spin in question, γ is the gyromagnetic ratio of the nucleus under study, B_0 is the strength of the magnetic field and $\hat{I}_z$ is an operator which denotes the nuclear spin angular momentum of spin I and in particular its z-component that interacts with the applied magnetic field.

Substituting equation 4.3 for the offset of the spin we get

$$\widehat{H}_{one\ spin} = -\gamma \Delta B \hat{I}_z = \Omega \hat{I}_z \qquad 4.6$$

During a period of free precession of a two-spin system, the offset and the J-coupling evolve, and the Hamiltonian is expressed as

$$\widehat{H}_{two\ spins} = v_{0,I}\hat{I}_z + v_{0,S}\hat{S}_z + J_{IS}\hat{I}_z\hat{S}_z \qquad 4.7$$

where S is the second spin, J_{IS} is the scalar coupling between the two spins and v_0 is the chemical shift in Hz ($v_0 = \frac{\omega_0}{2\pi} = \frac{-\gamma B_0}{2\pi}$), both in units of Hz.

In the rotating frame, equation 4.7 can be re-written by substituting for the offset of the proton, Ω_I, and the offset of the heteronuclear spin, Ω_S.

$$\widehat{H}_{two\ spins} = \Omega_I \hat{I}_z + \Omega_S \hat{S}_z + 2\pi J_{IS}\hat{I}_z\hat{S}_z \qquad 4.8$$

The scalar or J-coupling between two spins is a quantum mechanical property of two nuclei which have chemical bond connectivity and results in an observable doublet for each spin, rather than a singlet. Bonding electrons couple the nuclear spin to the coupled nucleus,

resulting in two microenvironments and as a result two slightly different chemical shifts. The chemical shift difference between the nuclei, in Hz, is the value of J.

The J-coupling is a key feature of many NMR experiments, where it is used to transfer magnetisation from one spin to another.

At each stage in the experiment, different kinds of magnetisation can be detected based on which axis the magnetisation of either spin is aligned on, e.g. $2\hat{I}_z\hat{S}_x$ indicates the proton magnetisation being on the z-axis and the heteronuclear spin magnetisation being on the x-axis. The different types of magnetisation are listed below.

Type of magnetisation	**Operators**
z-magnetisation on a single spin	$\hat{I}_z, \hat{S}_z$
in-phase magnetisation	$\hat{I}_x, \hat{I}_y, \hat{S}_x, \hat{S}_y$
anti-phase magnetisation	$2\hat{I}_x\hat{S}_z, 2\hat{I}_y\hat{S}_z, 2\hat{I}_z\hat{S}_x, 2\hat{I}_z\hat{S}_y$
multiple-quantum coherence	$2\hat{I}_x\hat{S}_x, 2\hat{I}_x\hat{S}_y, 2\hat{I}_y\hat{S}_x, 2\hat{I}_y\hat{S}_y$
non-equilibrium population	$2\hat{I}_z\hat{S}_z$

In product operator terms, J-coupling is used to change in-phase magnetisation to anti-phase magnetisation, during a time τ, e.g. $\hat{I}_x$ and $\hat{I}_y$ terms evolve into the anti-phase $2\hat{I}_x\hat{S}_z$ and $2\hat{I}_y\hat{S}_z$ terms. For a specific value of τ there can be complete conversion of the in-phase to the anti-phase magnetisation. This is a key feature of the Heteronuclear Single Quantum Coherence (HSQC) experiment which will be discussed in section 4.2.2.

We can determine the effect of pulses on magnetisation by following a simple rule.

$$\cos\theta \times original\ operator + \sin\theta \times new\ operator$$

The new operator can be determined by the diagrammatic representation of the x-, y- and z- axes representing the $\hat{I}_x$, $\hat{I}_y$ and $\hat{I}_z$ operators, respectively. Positive rotation about an axis changes the original operator into the new operator.

It is also important to note that rotating a magnetisation vector about its own axis, e.g. $\hat{I}_x$ about the x-axis, has no effect. Similarly, the above rule can be applied to the effect of chemical shift, but instead of cos θ and sin θ we now have cos Ωt and sin Ωt. In this case, Figure 4.4A illustrates which the new operator will be.

The spin echo pulse sequence is a building block of the Insensitive Nuclei Enhanced by Polarisation Transfer (INEPT) sequence, which in turn is a building block of many 2D NMR experiments, including the HSQC (Cavanagh *et al.*, 1998; Rule & Hitchens, 2006). The role of the INEPT pulse sequence is very important. The strength of NMR signals depends on the strength of the applied magnetic field and also on the gyromagnetic ratio of the nucleus under study. The stronger the applied magnetic field is, or the larger the gyromagnetic ratio of the nucleus, the greater the equilibrium magnetisation and therefore the higher the signal strength, i.e. higher sensitivity. The INEPT experiment is used to transfer magnetisation from a nucleus with a large gyromagnetic ratio, e.g. ^{1}H, to a nucleus with a smaller gyromagnetic ratio, e.g. ^{15}N (10 times as small as ^{1}H) or ^{13}C (4 times as small as ^{1}H). This has a dramatic effect on the enhancement of the NMR signals.

Rather than providing an in-detail description of the spin echo and INEPT pulse sequences, the use of these sequences will be illustrated in the next section which explains the basic, yet very important NMR experiment, the HSQC.

In this discussion spin I is always ^{1}H while spin S is the heteronuclear spin, i.e. ^{15}N or ^{13}C. The experiments always begin on the proton and then transfer magnetisation to the heteronuclear spin. This is because the gyromagnetic ratio of the proton is bigger. This leads to a stronger signal.

4.2.2 The HSQC pulse sequence

Figure 4.4 illustrates the HSQC pulse sequence .

At the start of the experiment, there is equilibrium magnetisation, which is aligned with the z-axis and is represented by the $\hat{I}_z$ operator. The first building block of the HSQC pulse sequence is the INEPT which achieves polarisation transfer. The first 90° x-pulse on the proton generates transverse magnetisation

$$\hat{I}_z \xrightarrow{90_I^x} -\hat{I}_y \qquad 4.9$$

During the 2τ period, the proton chemical shift is refocused due to the 180° pulse. The 180° pulse on the S causes the J-coupling (e.g. J_{NH}) to evolve. If

$$\tau = \frac{1}{4J} \qquad 4.10$$

this simplifies the magnetisation to

$$-2\hat{I}_x\hat{S}_z \qquad 4.11$$

The 90° y-pulse on I changes the magnetisation to $2\hat{I}_z\hat{S}_z$ and a 90° x-pulse on S at the same time changes this to $-2\hat{I}_z\hat{S}_y$, which can be summarised by

$$-2\hat{I}_x\hat{S}_z \xrightarrow[90_S^x]{90_I^y} -2\hat{I}_z\hat{S}_y \qquad 4.12$$

indicating longitudinal I magnetisation and transverse S magnetisation, i.e. magnetisation transfer has occurred, leading to a signal from the S, the magnitude of which is dependent upon the equilibrium magnetisation of I (usually proton).

During the t_1 period, the chemical shift of spin I is refocused (the J coupling between the two spins is also refocused) but the chemical shift of spin S evolves and thereby gets encoded.

The heteronuclear chemical shift evolves during the t_1 period to give

$$2\hat{I}_z\hat{S}_y \xrightarrow{\Omega_S t_1} 2\hat{I}_z\left[\hat{S}_y\cos(\Omega_S t_1) - \hat{S}_x\sin(\Omega_S t_1)\right] \qquad 4.13$$

$$= 2\hat{I}_z \hat{S}_y \cos(\Omega_S t_1) - 2\hat{I}_z \hat{S}_x \sin(\Omega_S t_1) \quad 4.14$$

The final building block in the HSQC is a reverse INEPT which achieves polarisation transfer back to the proton. An x-pulse is applied to both the proton and the heteronuclear spin.

$$2\hat{I}_z \hat{S}_y \cos(\Omega_S t_1) \xrightarrow{90_I^x} 2\hat{I}_y \hat{S}_z \cos(\Omega_S t_1) \quad 4.15$$

$$2\hat{I}_z \hat{S}_x \sin(\Omega_S t_1) \xrightarrow{90_S^x} -2\hat{I}_y \hat{S}_x \sin(\Omega_S t_1) \quad 4.16$$

Adding up the two terms gives

$$2\hat{I}_y \hat{S}_z \cos(\Omega_S t_1) - 2\hat{I}_y \hat{S}_x \sin(\Omega_S t_1) \quad 4.17$$

The $-2\hat{I}_y \hat{S}_x \sin(\Omega_S t_1)$ term is double quantum magnetisation which cannot be detected.

In the final step, the *J*-coupling evolves to give

$$2\hat{I}_y \hat{S}_z \cos(\Omega_S t_1) \xrightarrow{J} \hat{I}_x \cos(\Omega_S t_1) \quad 4.18$$

This step is important to convert anti-phase to in-phase magnetisation, which will result in two NMR peaks separated by the J-coupling.

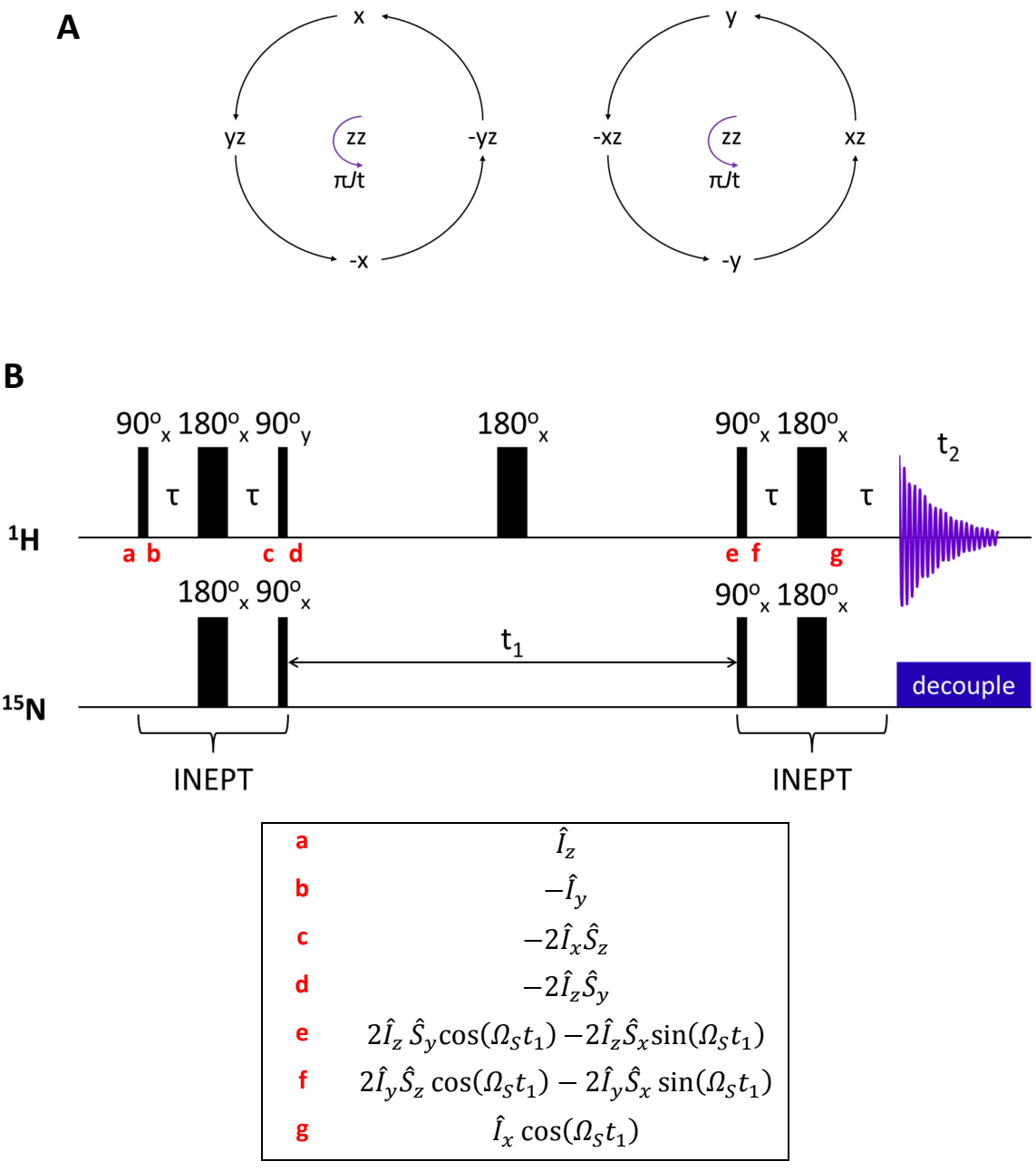
This allows decoupling to be applied during detection to give a single resonance line for each pair of I and S spins. The final signal is expressed as

$$S(t_1, t_2) = \cos(\Omega_S t_1) e^{-i\Omega_I t_2} \quad 4.19$$

A Fourier transformation is then used to give the final signal that is observed as a peak in a spectrum.

Finally, it is important that mention that NMR experiments on biological macromolecules are often recorded with water as the solvent. It is therefore imperative that we somehow remove the signal of the water protons, which are at a concentration of 111.2 M. This is done by including elements in the pulse sequence, such as the WATERGATE and water flip-back elements. An in-detail description of how these elements work can be found elsewhere (Piotto *et al.*, 1992).

Three-dimensional experiments use HSQC-based pulse sequences to transfer magnetisation from the ^{1}H to the ^{15}N to the ^{13}C. The pulse sequences of these experiments will not be discussed here. Instead, only the pathway of magnetisation transfer will be shown in each relevant section.



a	$\hat{I}_z$
b	$-\hat{I}_y$
c	$-2\hat{I}_x\hat{S}_z$
d	$-2\hat{I}_z\hat{S}_y$
e	$2\hat{I}_z\,\hat{S}_y\cos(\Omega_S t_1) - 2\hat{I}_z\hat{S}_x\sin(\Omega_S t_1)$
f	$2\hat{I}_y\hat{S}_z\cos(\Omega_S t_1) - 2\hat{I}_y\hat{S}_x\sin(\Omega_S t_1)$
g	$\hat{I}_x\cos(\Omega_S t_1)$

Figure 4.4 Product operator approach and the HSQC pulse sequence. (**A**) Diagrams for determination of the effect of *J*-coupling evolution on different types of magnetisation. (**B**) The HSQC pulse sequence. The product operators at each step (a-g) are noted in the key.

4.2.3 Other through-bond experiments

The HSQC is the basis of several other three-dimensional (3D) experiments which are important in the assignment of backbone NH chemical shifts. Four commonly recorded triple resonance experiments are the HNCA, the HNCACB, the HN(CO)CA and the HN(CO)CACB (Cavanagh *et al.*, 1998; Rule & Hitchens, 2006). These give correlations between the amide ^{1}H and ^{15}N and the backbone $^{13}C_{\alpha}$ and $^{13}C_{\beta}$. Figure 4.11 shows how the HSQC is the basis of these experiments and how the peaks are further resolved in the ^{13}C dimension. Figure 4.10 summarises the pathway of magnetisation transfer in each of these experiments.

An additional through-bond experiment which is useful in assigning the side-chain chemical shifts is the HCCH-TOCSY. This provides a total correlation spectrum which shows which protons are connected via a chemical bond. This can be linked to the ^{13}C-HSQC spectrum which can help the assignment process.

4.2.4 The NOE

The Nuclear Overhauser Effect (NOE) is a property of dipolar relaxation between two nuclear spins. This means that two atoms which may be far away in the protein primary sequence, but up to 5 Å close in space, will give rise to an NMR signal as a result of successful magnetisation transfer through space from one spin to the other. The strength of this NOE signal is inversely proportional to r^{6}, where r is the distance between the two nuclei. The intensities of peaks in Nuclear Overhauser Effect Spectroscopy (NOESY) spectra are used as distance restraints in structure calculations.

4.3 NMR studies on PRK1 HR1c

4.3.1 2D-NOESY and 2D-TOCSY of unlabelled PRK1 HR1c

Two-dimensional (2D) NMR experiments such as the 2D NOESY and 2D TOCSY (Cavanagh *et al.*, 1998; Rule & Hitchens, 2006) are a rapid way of obtaining preliminary information on the secondary structure of proteins, as well as information on their overall fold. The 2D NOESY experiment uses the NOE to report on protons which are close in space. Depending on the secondary structure of the protein in question, a different region of the spectrum is populated with peaks. For example, in an α-helix the amide protons of sequential residues are arranged close to each other. This gives rise to strong NOEs between these protons at chemical shifts of 7-9 ppm. In contrast, β-sheet amide protons are located close to the H_α of the preceding residue and the resulting NOEs populate an alternative region of the NOESY spectrum. The 2D TOCSY experiment transfers magnetisation through chemical bonds via the ^3J ^{1}H-^{1}H couplings and shows cross-peaks between an amide proton and side chain protons (H_α, H_β, H_γ, etc.). Both of these spectra can also provide information on the feasibility of pursuing structural work by NMR. If the peaks are uniformly strong and sharp, then the protein is likely to be monomeric (or a small dimer). This also practically means that the double and triple resonance experiments which are crucial in structural studies will have good resolution and minimal peak overlap. Conversely, if the peaks are broadened, then there is indication of exchange processes such as protein oligomerisation, which would be a significant impediment when recording other NMR experiments as the resolution of recorded spectra would be poor.

The starting point of this work was to record ^{1}H homonuclear 2D NOESY and 2D TOCSY spectra on an unlabelled PRK1 HR1c NMR sample (Figure 4.5). Uniformly strong and sharp peaks were observed which suggested that other NMR experiments to be recorded on labelled protein would also be of comparable quality. In addition, NOE peaks were observed indicative of backbone NH-NH NOEs in an α-helical protein. This indicates that the protein is structured and in particular α-helical and it is in agreement with the CD results recorded previously (Chapter 3). Furthermore, the absence of broadened peaks is indicative of the

lack of oligomerisation or non-specific protein aggregation, which is consistent with the AUC analysis performed on this protein (Chapter 3).

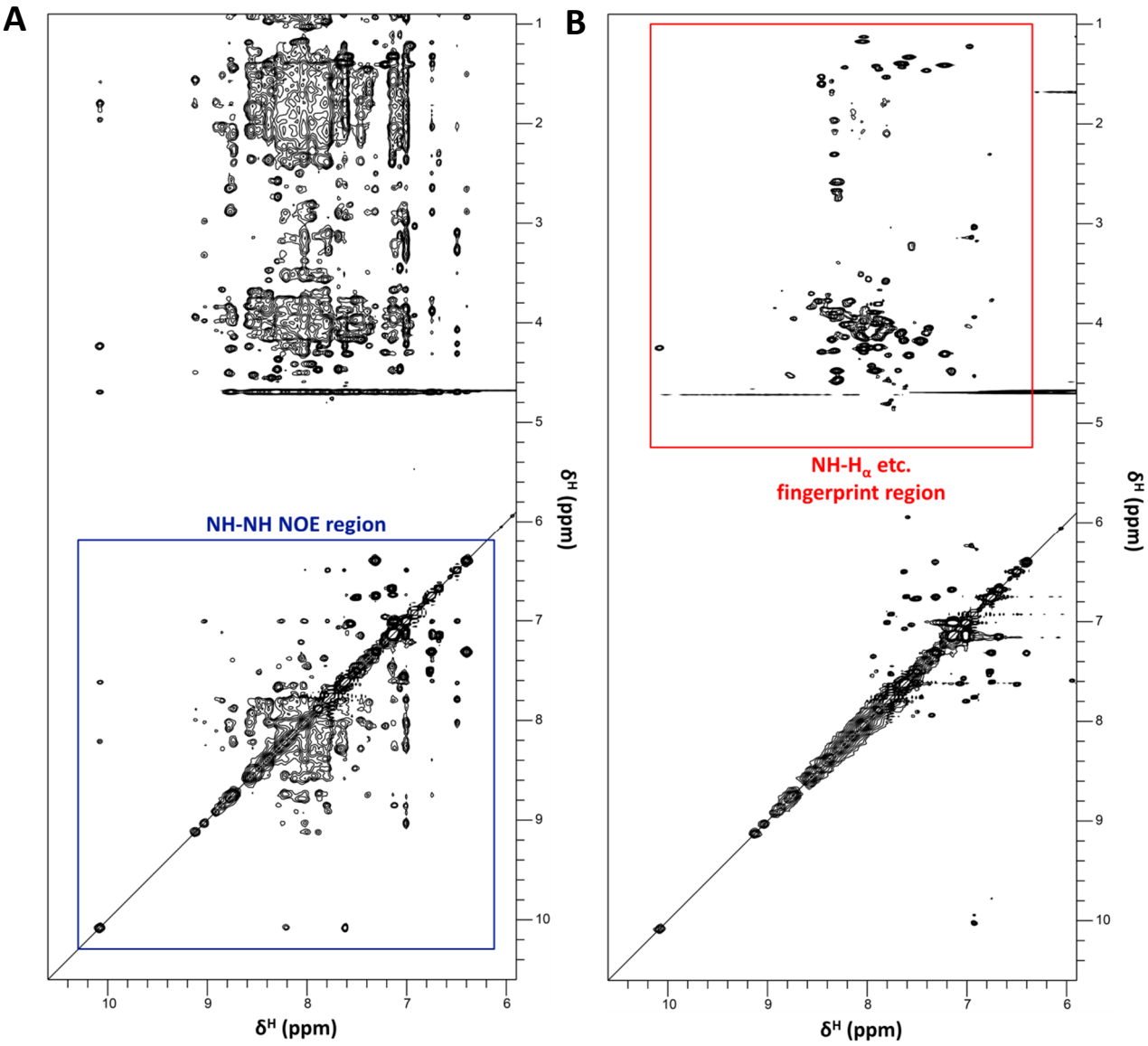


Figure 4.5 2D NOESY and 2D TOCSY spectra of PRK1 HR1c. NMR spectra of 1.6 mM PRK1 HR1c in 20 mM sodium phosphate pH 7.3, 150 mM NaCl and 10% D_2O were recorded on a Bruker DRX500 at 25 ^{o}C. The amide region which contains NH-NH NOEs is highlighted in the 2D NOESY while the NH-H_α etc. fingerprint region is highlighted in the 2D TOCSY.

4.3.2 Purification of ^{15}N-labelled NMR sample

To record the double and triple resonance experiments on backbone amides and also side chain groups (CH) needed for the structural studies of HR1c, the protein would have to include ^{15}N and ^{13}C nuclei, which are the spin active forms of N and C but are of very low natural abundance (0.368% for ^{15}N and 1.07% for ^{13}C (Rosman & Taylor, 1998)). To isotopically enrich the protein with ^{15}N and ^{13}C, $^{15}NH_4Cl$ and ^{13}C glucose are supplied to a bacterial culture as the only sources of nitrogen and carbon, respectively. These cultures can be grown in either MOPS or M9 minimal medium. Sometimes the expression can be significantly lower in one of the two media, so the expression of the protein in both of these media, and how it compares to expression in 2TY medium, was assessed by small scale expression. The protein expressed equally well in both media and M9 was chosen as the minimal medium in which to express the labelled protein (Figure 4.6). The protein was purified as shown previously (Chapter 3) (Figure 4.7) and 10% D_2O was added giving an NMR sample with final concentration of 1.6 mM.

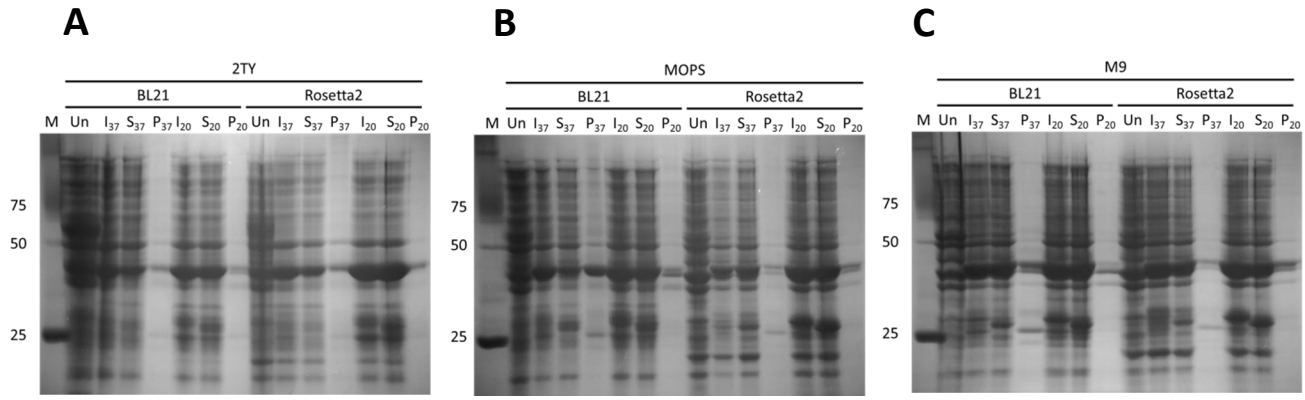


Figure 4.6 SDS-PAGE Analysis of ^{15}N PRK1 HR1c expression trials in minimal medium. Solubility of fusion protein was assessed by identifying band of approximate molecular weight in soluble or pellet fractions. Un for Uninduced sample, In for Induced (Total), S for Induced (Soluble) and P for Induced (Pellet). 20 and 37 denote the temperature at which the expression was carried out at. Laemmli gels were 12%. (**A**) 2TY (**B**) MOPS (**C**) M9

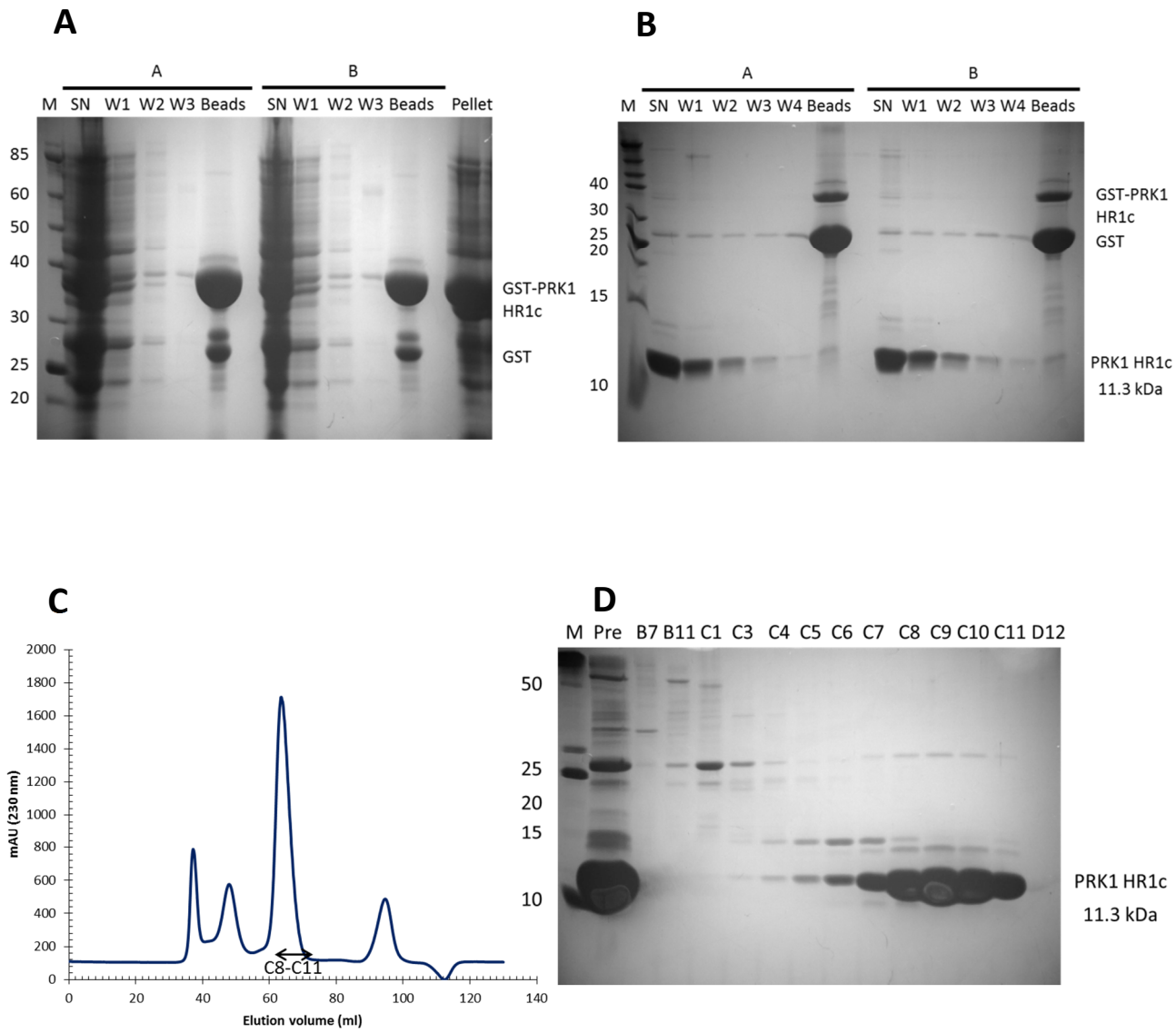


Figure 4.7 Purification of ^{15}N PRK1 HR1c and corresponding SDS-PAGE analysis. (**A**) GST-PRK1 HR1c immobilised on glutathione agarose beads. Gel was 12%. SN for supernatant, W1-3 for the 3 washes, Beads for 50% bead slurry and Pellet for insoluble fraction. (**B**) PRK1 HR1c washes post-cleavage. Gel was 18%. SN for supernatant and W1-4 for the 4 washes post-cleavage. (**C**) Gel filtration chromatogram of PRK1 HR1b on an S75 column. Blue line indicates A_{230}. (**D**) SDS-PAGE analysis of chromatogram in (**C**) on an 18% Laemmli gel. 'Pre' for sample loaded on the column. The double arrows indicate the fractions pooled.

4.3.3 ^{15}N-HSQC and ^{13}C-HSQC experiments

The fundamental NMR experiment recorded in protein studies is the HSQC. The ^{15}N-HSQC and the ^{13}C-HSQC provide through-bond correlations between proton and nitrogen or proton and carbon chemical shifts, respectively. Typically, the ^{15}N-HSQC has a much smaller number of expected peaks than the ^{13}C-HSQC does and therefore there is less overlap. The ^{15}N-HSQC spectrum comprises peaks corresponding to all NH groups in the protein. These peaks include backbone NH groups (prolines lack NH groups) and side-chain NH_2 groups of asparagine and glutamine residues. There are two peaks for each of the side-chain NH_2 groups, one for each of the protons. These have different ^{1}H chemical shifts but an identical ^{15}N chemical shift and they are located in the top right region of the spectrum, so they are very easily identified. Arginine side-chain NH groups are often visible but may be weak as a result of exchange broadening. Instead of appearing at very small ^{15}N chemical shifts, they are instead aliased to larger ^{15}N chemical shifts. Additionally, while the side-chain NH group of tryptophan is observable, the NH group of the histidine and the NH_3^+ group of the lysine side-chains are often not. The N-terminal NH_3^+ group was not visible due to broadening caused by exchange with the water solvent.

The purified PRK1 HR1c protein has 104 amino acids; 7 amino acids, GPLGSHM come from the vector sequence and include a remainder of the 3C protease cleavage site (GPL) as well as residues encoded by the restriction sites used during cloning (GSHM).The remaining 97 amino acids correspond to the HR1c domain (residues 201-297 of the full length PRK1). A ^{15}N-HSQC spectrum was recorded on the purified sample (Figure 4.8). HR1c contains 5 proline residues, so the expected number of peaks corresponding to backbone amide groups is 98. The ^{15}N-HSQC spectrum was expected to contain a total of 108 peaks; 98 backbone NH groups and 10 corresponding to the 5 pairs of asparagine and glutamine side-chain NH_2 groups. Instead, only 95 peaks were observed, suggesting that there may be some peak overlap. Nevertheless, the spectrum is well-dispersed (δ^H = 7.0-10.4 ppm and δ^N = 104-126 ppm) which suggests that the protein is structured and also that pursuing the structure is amenable by NMR. The sharpness of the peaks is also consistent with the protein being monomeric. The 10 asparagine and glutamine NH_2 peaks were easily identified for reasons

explained above at δ^H = 6.4-7.6 ppm and δ^N = 109-114 ppm. 9 out of 10 arginine side-chain ($H_\varepsilon N_\varepsilon$) peaks were also observed.

To record experiments that would enable the assignment of backbone and side-chain chemical shifts, ^{15}N, ^{13}C-labelled protein was purified in the same way as the ^{15}N-labelled protein and concentrated to yield a sample with final concentration of 1.2 mM. The recorded ^{13}C-HSQC experiment can be seen in Figure 4.9. This spectrum contains a peak for each CH group in the protein and it is noticeably more crowded than the ^{15}N-HSQC. The important regions of the spectrum have been annotated to show the different categories of CH groups. This spectrum was used to aid side-chain assignments in conjunction with the HCCH-TOCSY and ^{13}C-separated NOESY experiments (Section 4.3.6).

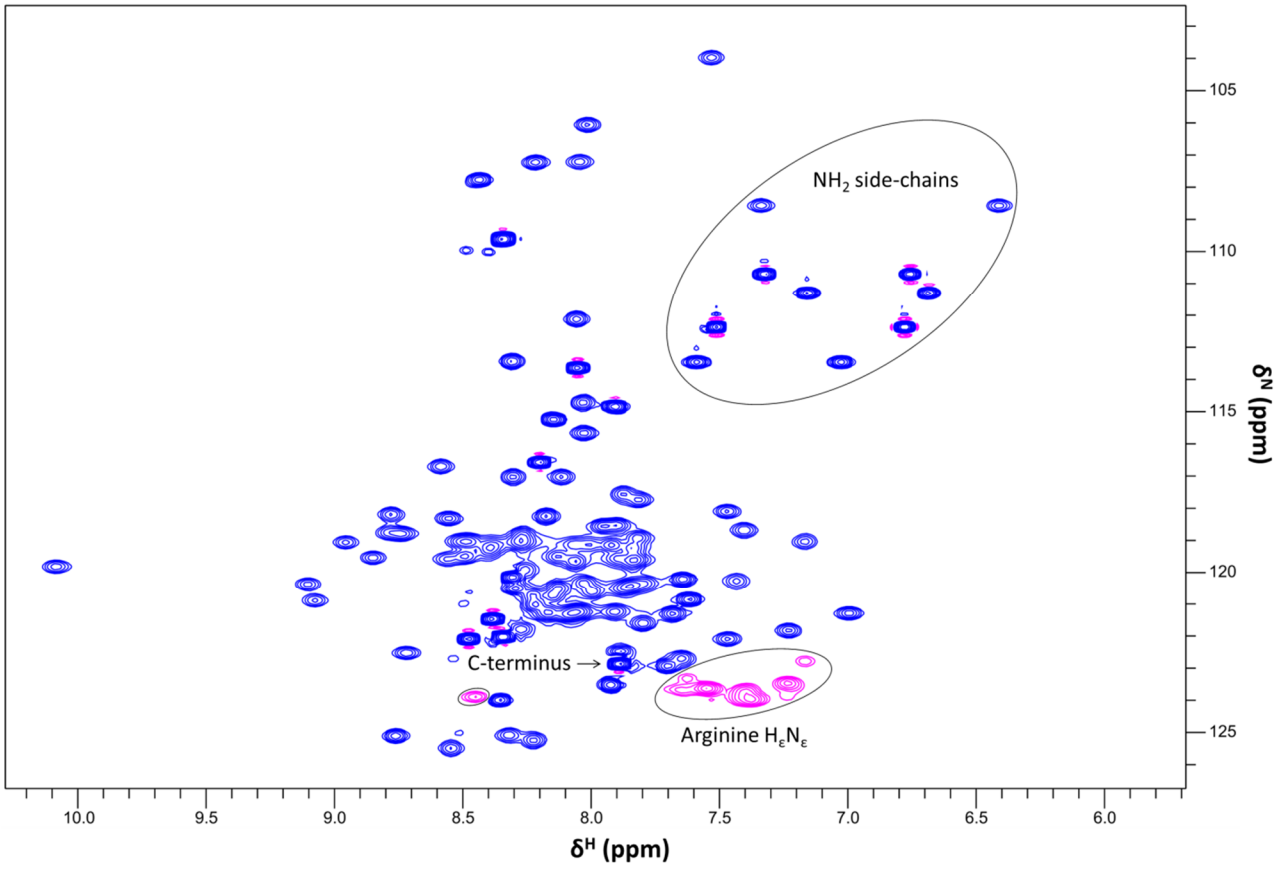


Figure 4.8 ^{15}N-HSQC spectrum of PRK1 HR1c. A ^{15}N-HSQC of 1.6 mM PRK1 HR1c in 20 mM sodium phosphate pH 7.3, 150 mM NaCl and 10% D_2O was recorded on a Bruker DRX500 at 25 ^oC. Asparagine and glutamine side-chain NH_2 groups, arginine side-chain $N_\varepsilon H_\varepsilon$ groups (aliased in the N dimension and negative, coloured in pink) and the C-terminus are highlighted.

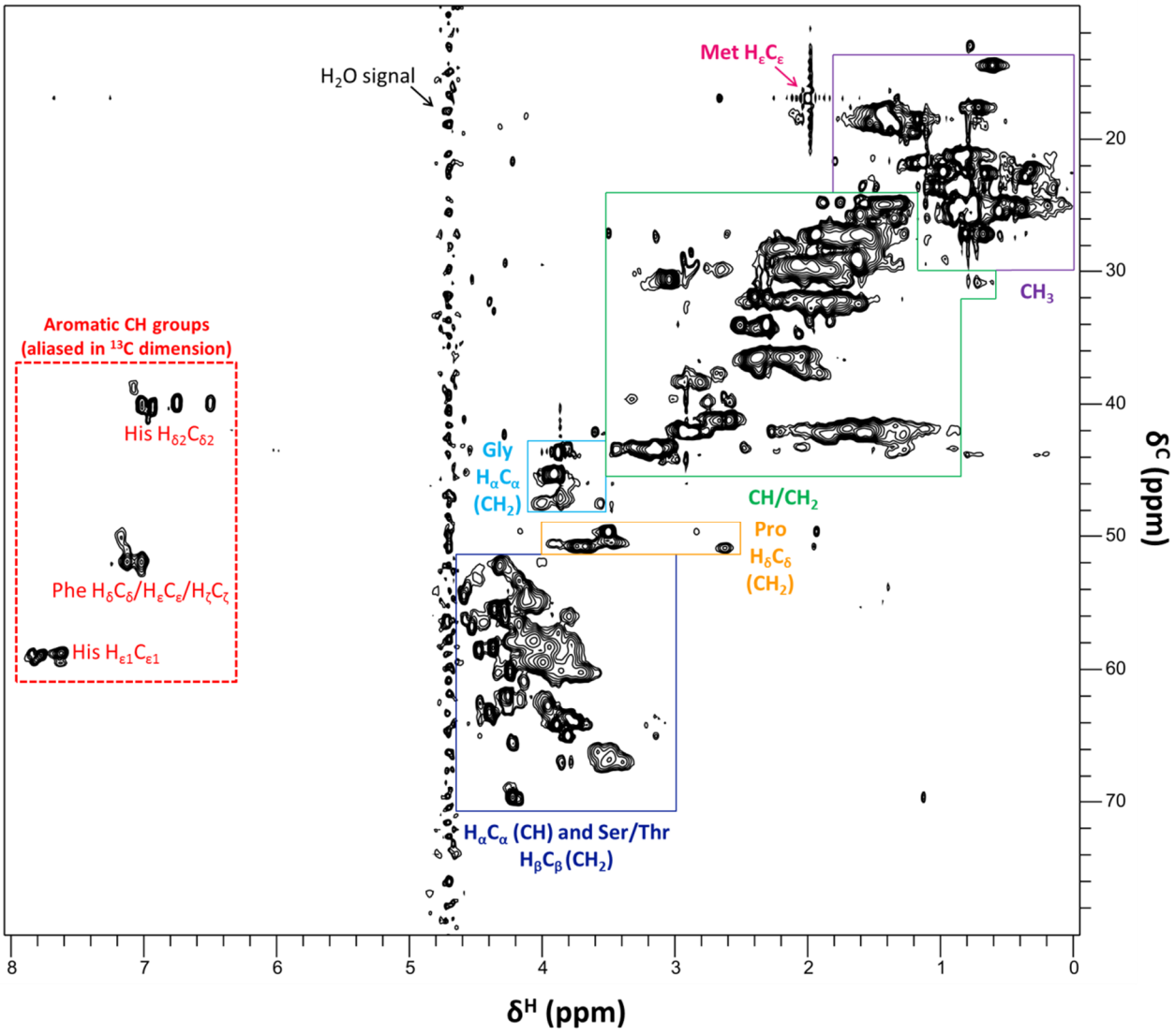


Figure 4.9 ^{13}C-HSQC spectrum of PRK1 HR1c. A ^{13}C HSQC of 1.2 mM PRK1 HR1c in 20 mM sodium phosphate pH 7.3, 150 mM NaCl and 10% D_2O was recorded on a Bruker DRX500 at 25 ^{o}C. The various aliphatic and aromatic CH group regions are highlighted.

4.3.4 Three-dimensional experiments for backbone NH group assignment

Assigning the chemical shifts of backbone and side-chain protons is an obligatory requirement in protein structural studies. Dihedral angle restraints can be derived from the backbone chemical shifts, while the assignments can be used to derive proton distance restraints from NOESY experiments. Together these restraints form the basis of structure calculations. The assignment process begins by assigning the backbone NH shifts using a series of triple resonance experiments. The side-chains are then assigned by recording additional three-dimensional experiments.

To assign the backbone NH shifts, four three-dimensional experiments were recorded, the HNCA (Kay *et al.*, 1990), the HNCACB, the HN(CO)CA and the HN(CO)CACB (Figure 4.10). The HNCA gives an intra-residue correlation between the ^{1}H and ^{15}N shifts of the backbone amide and the ^{13}C shift of the alpha carbon atom, C_α. In addition, some magnetisation can be transferred by the two-bond coupling to the C_α of the preceding residue, giving rise to an additional, but weaker peak in the spectrum which can aid the assignment process. The HNCACB experiment provides the same correlations that the HNCA does, but in addition it gives correlations between the ^{1}H and ^{15}N shifts and both the intra-residue C_β and also that of the preceding residue. The HN(CO)CA experiment gives a single correlation between the backbone ^{1}H and ^{15}N and the C_α of the preceding residue, and is therefore useful when comparing with the HNCA experiment to identify the C_α shift of the preceding residue unambiguously. Finally, the HN(CO)CACB experiment complements the HN(CO)CA, providing the same information but in addition a correlation between the ^{1}H and ^{15}N of residue i and the C_α and C_β of residue i_{-1}, and it is used in conjunction with the HNCACB. Together, these four experiments establish connectivity between the residues and in this way, we can "walk the backbone", which allows assignment of the backbone chemical shifts.

Sequential assignment was aided by knowledge of the ^{13}C chemical shifts of all amino acids found in databases such as the Biological Magnetic Resonance Bank. For example, the $^{13}C_\alpha$ chemical shift of alanine is approximately expected to be around 51 ppm, while the $^{13}C_\beta$ chemical shift is expected to be close to a characteristic shift of 19 ppm. Starting points in sequential assignment included residues such as leucine and aspartate with characteristic

$^{13}C_\beta$ (~40 ppm) as well as glycine, which has no C_β, so only two peaks appear in the HNCACB spectrum and one in the HN(CO)CACB spectrum. Importantly, the chemical shifts can be affected by different secondary structures. For example, in an α-helical context, carbon chemical shifts are shifted up-field compared to those carbons in a random coil. This is a result of the carbonyl groups in an α-helix having an effect on the electronegative properties of these atoms (Spera & Bax, 1991). Likewise, side-chain proton chemical shifts are shifted down-field compared to those protons in a random coil.

As prolines lack backbone NH groups, they don't give rise to any peaks in the HNCA and HNCACB experiments. However, in the HN(CO)CA and HN(CO)CACB experiments magnetisation can be transferred from the NH of the succeeding-to-the-proline residue, so the C_α and C_β shifts of prolines can be assigned in this way. This becomes very helpful for predicting dihedral angle restraints of prolines but also assists side-chain assignments (Section 4.3.6).

The assignment process consisted of three steps (Figure 4.11). Firstly, a peak in the ^{15}N-HSQC was chosen and used to navigate to the corresponding ^{15}N strip in the HNCA/HNCACB/HN(CO)CA/HN(CO)CACB spectra. The chemical shift of the C_α of the i_{-1} residue was distinguished from the chemical shift of the C_α of the i residue by comparing the HNCA and the HN(CO)CA spectra. Secondly, the C_α chemical shift of the i_{-1} residue was used to search for a ^{15}N strip in the HNCA spectrum which contains a peak of matching chemical shift. Thirdly, there was navigation back to the ^{15}N-HSQC to find the corresponding peak of the i_{-1} residue. In this way, a sequential link was established between the two amide peaks in the HSQC. "Forward walking" of the backbone was also performed. A C_α peak (residue i) was selected in the HNCA strip and the corresponding HN(CO)CA strip of residue i_{+1} was located by finding a peak with a matching chemical shift. The C_β chemical shift was used alongside C_α for shift matching by comparing HNCACB and HN(CO)CACB experiments. There was subsequent navigation in the ^{15}N-HSQC to link the i and i_{+1} residues.

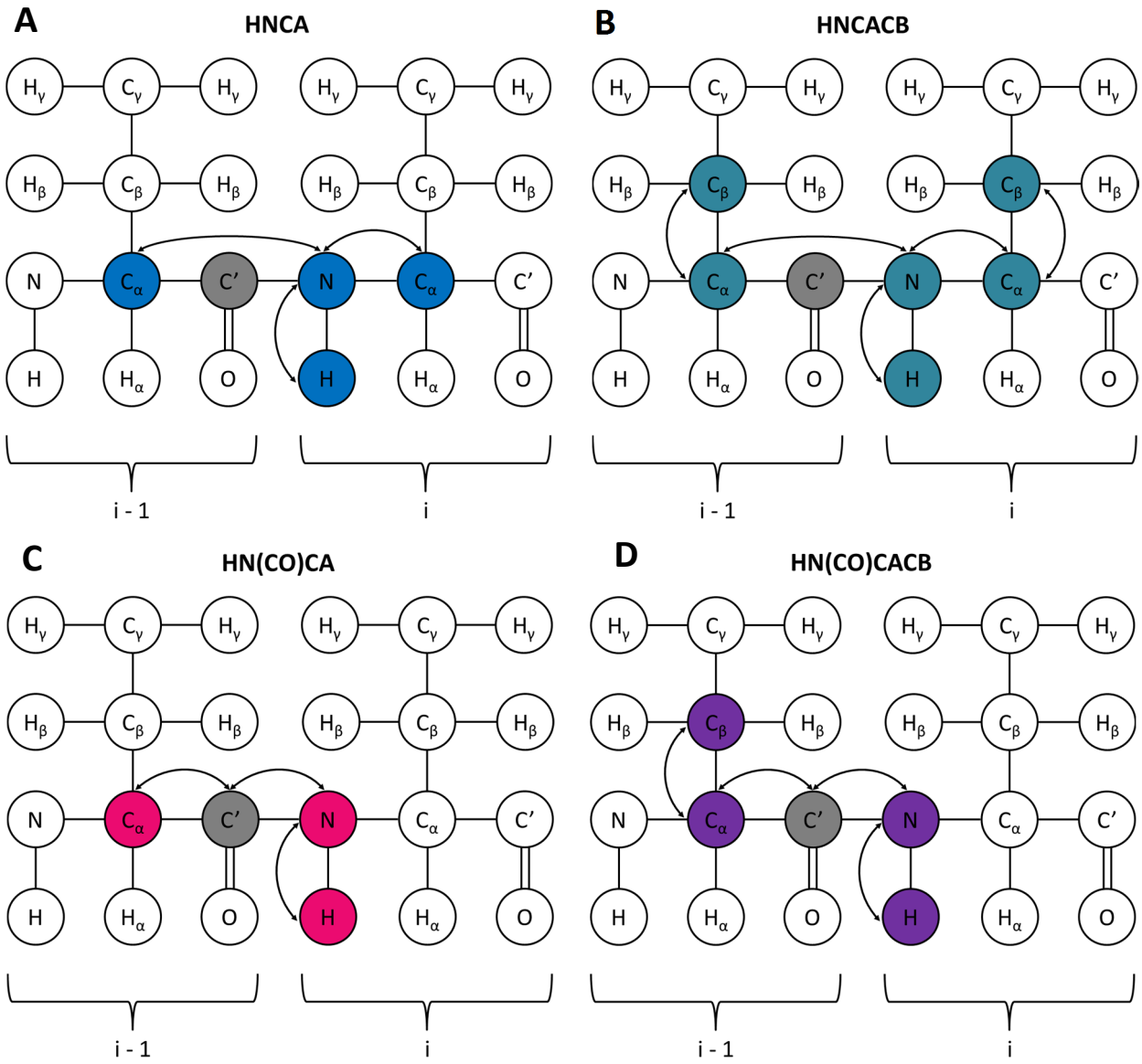


Figure 4.10 Triple resonance experiment magnetisation transfer pathways. Magnetisation transfer pathways of (**A**) HNCA (**B**) HNCACB (**C**) HN(CO)CA and (**D**) HN(CO)CACB. The chemical shift of the coloured nuclei is recorded in these experiments. The arrows indicate the through-bond magnetisation transfer using one- or two-bond *J*-couplings. The carbonyl carbon is used in the magnetisation transfer but its chemical shift is refocused.

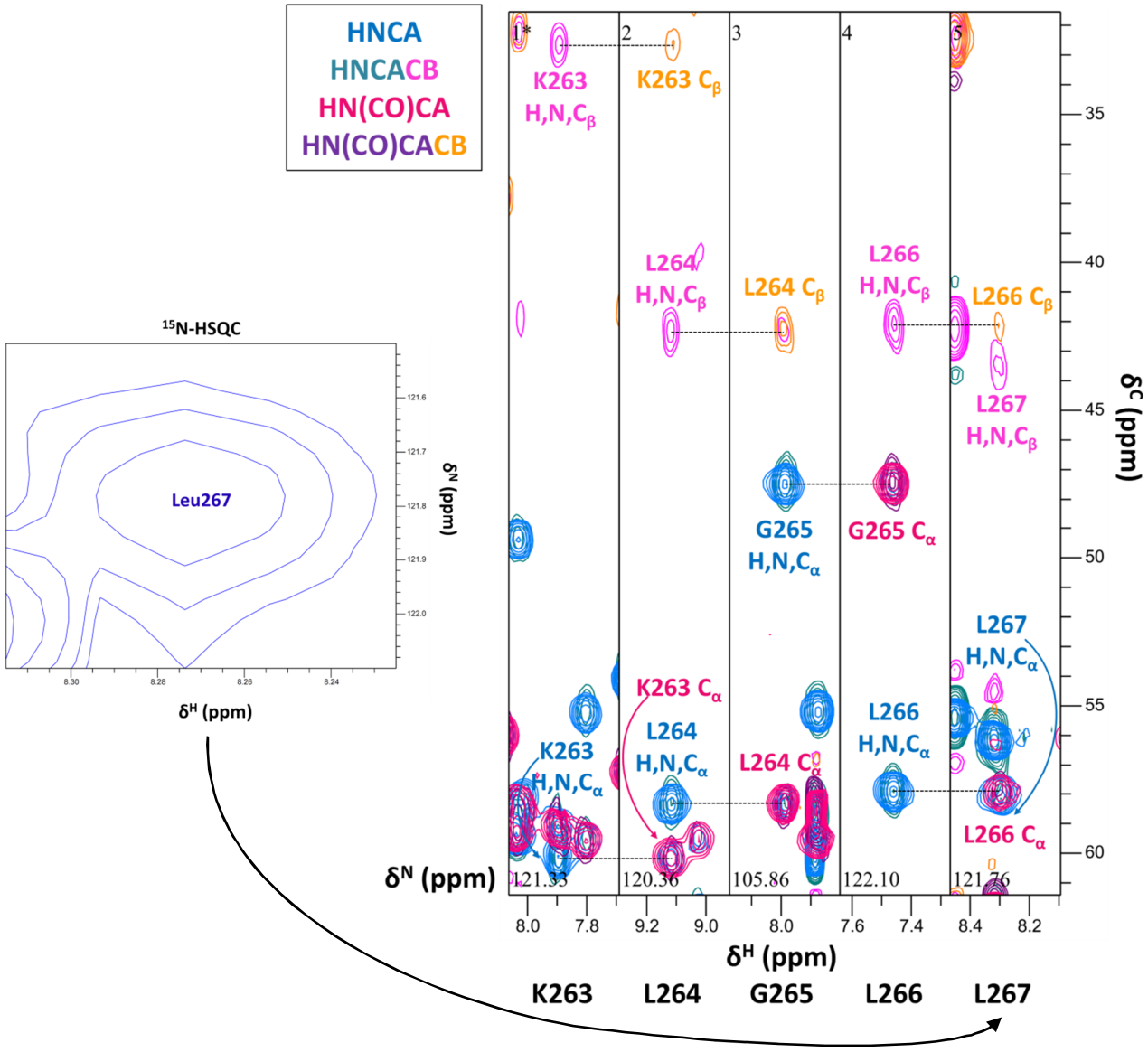


Figure 4.11 Backbone NH assignment using triple resonance experiments. A peak in the ^{15}N-HSQC was used to navigate to the corresponding ^{15}N strip in the HNCA/HNCACB/HN(CO)CA/HN(CO)CACB spectra. The chemical shift of the C_α of the i_{-1} residue was distinguished from the chemical shift of the C_α of the i residue by comparing the HNCA and the HN(CO)CA spectra. Secondly, the C_α chemical shift of the i_{-1} residue was then used to locate an ^{15}N strip in the HNCA spectrum which contains a peak of matching chemical shift. Thirdly, there was navigation back to the ^{15}N-HSQC to find the corresponding peak of the i_{-1} residue. In this way, a sequential link was established between the two amide peaks in the HSQC. "Forward walking" of the backbone was also possible. A C_α peak (residue i) was selected in the HNCA strip and the corresponding HN(CO)CA strip of residue i_{+1} was located by finding a peak with a matching chemical shift. There was subsequent navigation in the ^{15}N-HSQC to link the i and i_{+1} residues. The C_β chemical shifts were determined in the same way by comparing the HNCACB and HN(CO)CACB experiments. The figure shows the ^{15}N strips of five sequential residues. HNCA in blue, HNCACB in teal and pink, HN(CO)CA in magenta and HN(CO)CACB in purple and orange.

4.3.5 ^{15}N-separated TOCSY and ^{15}N-separated NOESY for backbone assignment

To aid the assignment of backbone amide chemical shifts, two further experiments were recorded on the ^{15}N PRK1 HR1c sample, the ^{15}N-separated TOCSY and the ^{15}N-separated NOESY. These are both three-dimensional experiments based on the HSQC and homonuclear ^{1}H-^{1}H couplings. They provide correlations between the backbone NH atoms and the side-chain protons (Figures 4.12A-B). In theory, correlations between all side-chain protons should be visible in the ^{15}N-separated TOCSY spectrum. However, this was often not the case, as a result of the small *J*-coupling (4-6 Hz) in helical proteins, which results in inefficient magnetisation transfer. These two spectra were used in parallel to help distinguish between intra-residue and inter-residue NOEs, as the ^{15}N-separated TOCSY spectrum correlates intra-residue protons (Figure 4.12C).

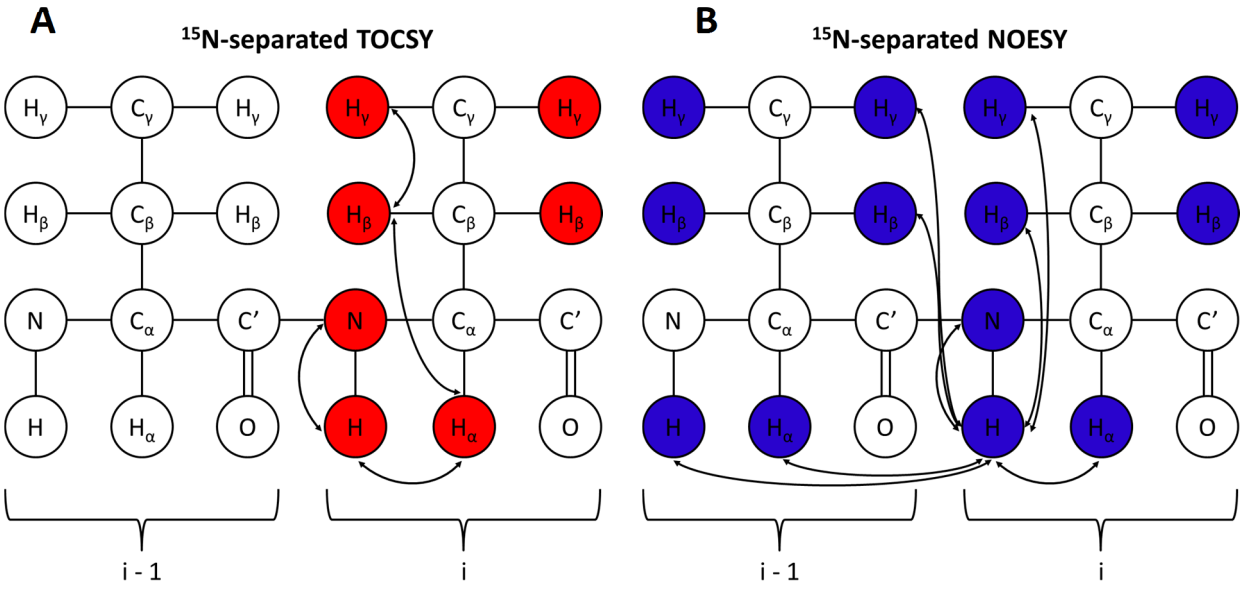


Figure 4.12 ^{15}N-separated NOESY and ^{15}N-separated TOCSY experiments. (**A**) Magnetisation transfer pathway of ^{15}N-separated TOCSY. In this experiment, magnetisation is transferred through bond using the *J*-coupling. (**B**) Magnetisation transfer pathway of ^{15}N-separated NOESY. In this experiment, magnetisation is transferred through space using the NOE. The arrows indicate the involved nuclei.

The ^{15}N-separated NOESY spectrum contains cross peaks for the NH-NH NOEs present in α-helices. These NOEs can be used as an alternative to the triple resonance experiments to assign the NH peaks in the ^{15}N-HSQC (Figure 4.13A). A peak was located in the ^{15}N-HSQC and used to navigate to the corresponding ^{15}N strip in the ^{15}N-separated NOESY/^{15}N-separated TOCSY window. The NH-NH NOE was selected and used to navigate to the orthogonal plane of these spectra and identify a return NOE peak. A return peak was located and used to navigate to a new strip in the ^{15}N-separated NOESY/^{15}N-separated TOCSY window and finally back to the ^{15}N-HSQC to establish connectivity between two peaks. This method was used to confirm the backbone assignments from the triple resonance experiments but to also assign the 10 backbone amides that could not be assigned using that set of experiments, e.g. Glu291 and Ala244. Since the HR1c domain is predicted to be α-helical, NOEs between Ha_i to $NH_{i,i+3}$ protons were also observed. Figure 4.13B shows an example of aligned ^{15}N strips of five consecutive residues and the NH-NH NOEs connecting them.

All backbone amides were assigned using a combination of the two approaches described above with the exception of a His amide found in the N-terminus of the protein and corresponding to vector sequence.

The ^{15}N-separated NOESY spectrum provided the first out of two sets of distance restraints for structure calculations.

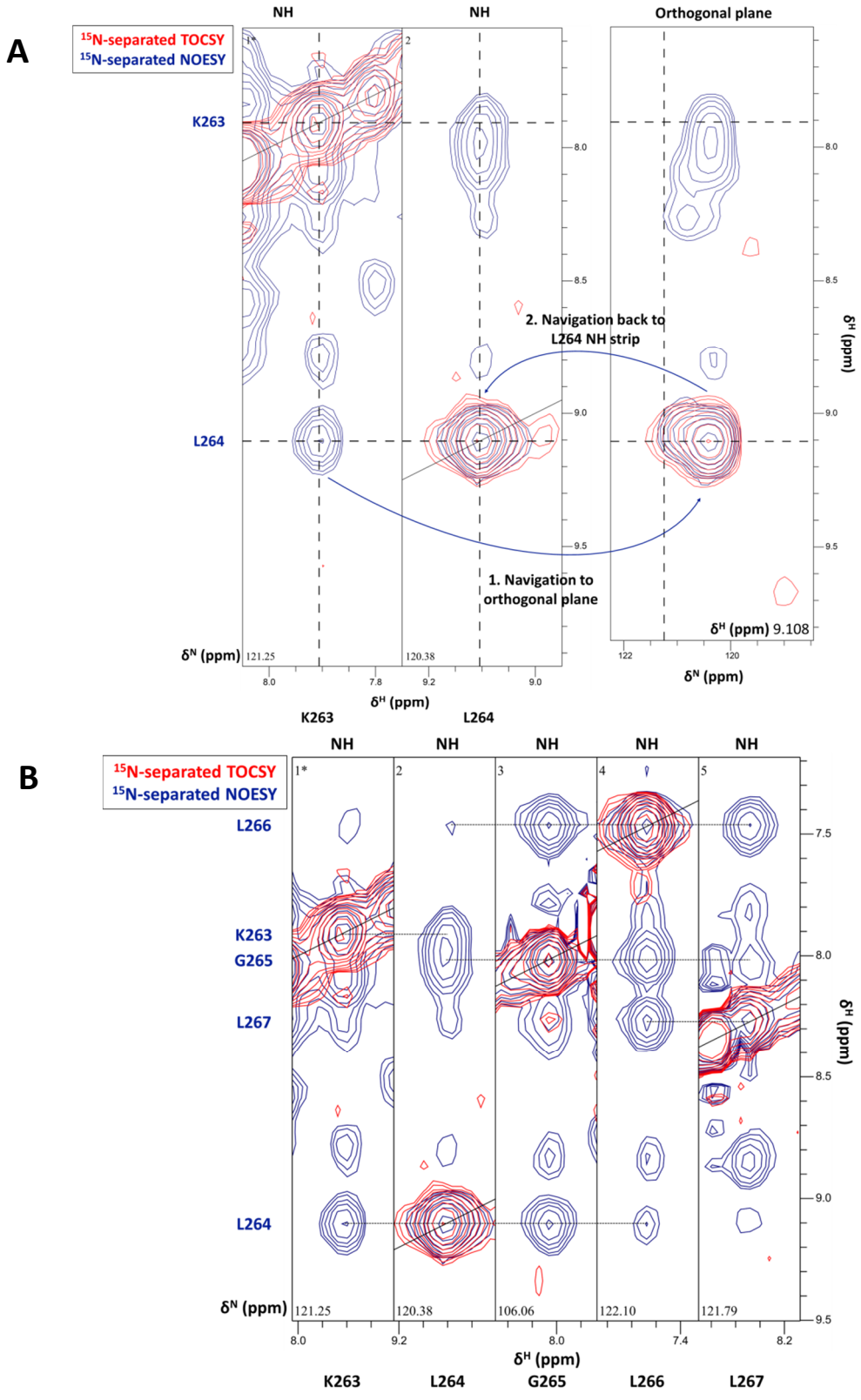


Figure 4.13 ^{15}N-separated NOESY spectrum of PRK1 HR1c. A ^{15}N-separated NOESY of 1.6 mM PRK1 HR1c in 20 mM sodium phosphate pH 7.3, 150 mM NaCl and 10% D_2O was recorded on a Bruker DRX500 at 25 ^{o}C. ^{15}N-separated NOESY in navy and ^{15}N-separated TOCSY in red. (**A**) An NOE peak was used to navigate to the orthogonal plane of the spectrum to locate a matching NOE pattern and then navigate back to a new ^{15}N strip. (**B**) Aligned ^{15}N strips of five consecutive residues. The NH-NH NOEs between them are linked with dotted lines.

4.3.6 HCCH-TOCSY and ^{13}C-separated NOESY

The HCCH-TOCSY was recorded on the double-labelled sample. Similar to the ^{15}N-separated TOCSY, it provides through-bond correlations between side-chain protons bound to a ^{13}C, so it is used to assign side-chain ^{1}H and ^{13}C chemical shifts (Figure 4.14A). Peaks picked in this spectrum were used to pick the ^{13}C HSQC spectrum. Conversely, unassigned regions in the ^{13}C HSQC spectrum were also used to navigate to the HCCH-TOCSY spectrum to help assign missing resonances.

Similarly, the ^{13}C-separated NOESY experiment (Figure 4.14B) was used to assign NOEs in side-chain protons. Through comparison with the HCCH-TOCSY, intra-residue and inter-residue NOEs were able to be distinguished. This experiment was particularly useful when assigning intra-residue protons, as the HCCH-TOCSY is sometimes significantly overlapped. The ^{13}C-separated NOESY provided the second set of distance restraints used in structure calculations.

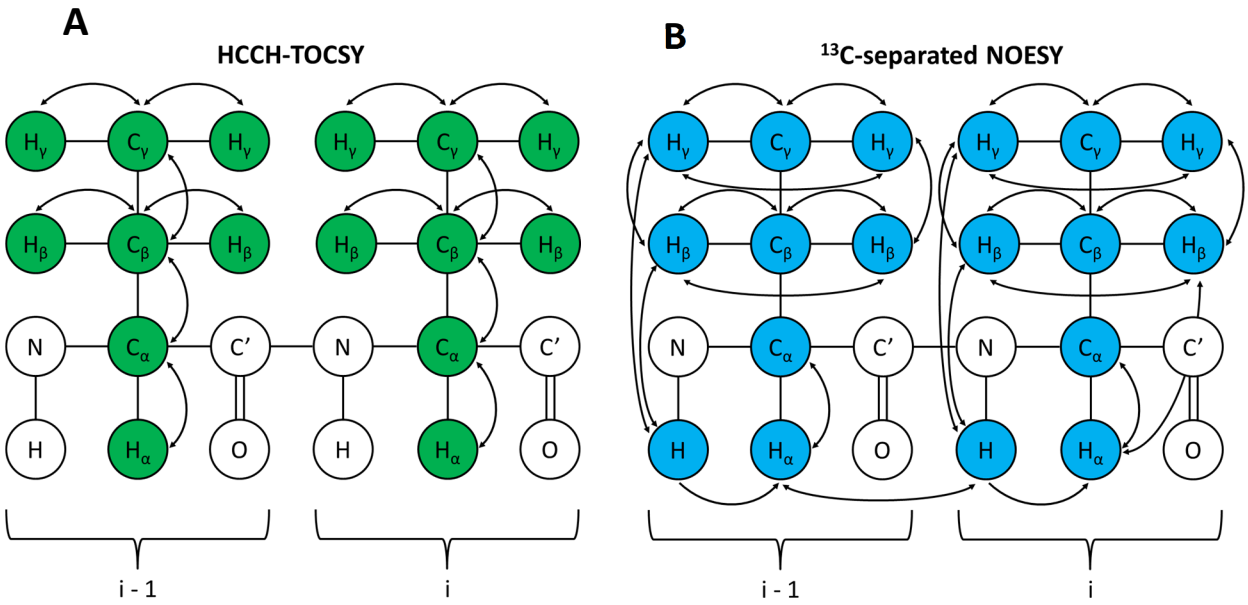


Figure 4.14 Magnetisation transfer pathways of HCCH-TOCSY and ^{13}C-separated NOESY. (**A**) Magnetisation transfer pathway of HCCH-TOCSY. In this experiment, magnetisation is transferred through bond using the *J*-coupling. (**B**) Magnetisation transfer pathway of ^{13}C-separated NOESY. In this experiment, magnetisation is transferred through space using the NOE. The arrows indicate the involved nuclei.

4.3.7 Assignment of aliphatic side-chains

The ^{1}H, ^{15}N, $^{13}C_{\alpha}$ cross-peak of a residue, e.g. Asp281 (Figure 4.15) was located in the HNCA spectrum. It was then used to navigate to the corresponding CH strip in the HCCH-TOCSY spectrum. The same procedure was repeated for the ^{1}H, ^{15}N, $^{13}C_{\beta}$ cross-peak in the HNCACB spectrum. Peaks in the HCCH-TOCSY and corresponding to magnetisation transferred from the $^{1}H_{\alpha}$ to the $^{1}H_{\beta}$ were selected and assigned. The ^{15}N-separated TOCSY spectrum was used to aid the assignment procedure, particularly if the H_{α} and other side-chain protons had already been assigned in that spectrum. Figure 4.15 shows an example of the assignment procedure of a short aliphatic side-chain, Asp281.

If the amino acid had additional CH groups in the side-chain, the TOCSY peaks were selected and were used to navigate to the orthogonal plane of the spectrum (Figure 4.16). From that view, a TOCSY pattern matching with the $^{1}H_{\alpha}$ and the $^{1}H_{\beta}$ was found at a ^{13}C chemical shift characteristic of that particular CH group, e.g. Leu271 $H_{\gamma}C_{\gamma}$. Then there was navigation from the orthogonal plane back to the normal plane at the $H_{\gamma}C_{\gamma}$ strip. The TOCSY and NOESY peaks from that proton, e.g. H_{γ}, to the other intra-residue protons, e.g. H_{α} and H_{β}, were then assigned. The process was then continued until >95% ^{1}H and ^{13}C chemical shifts were assigned.

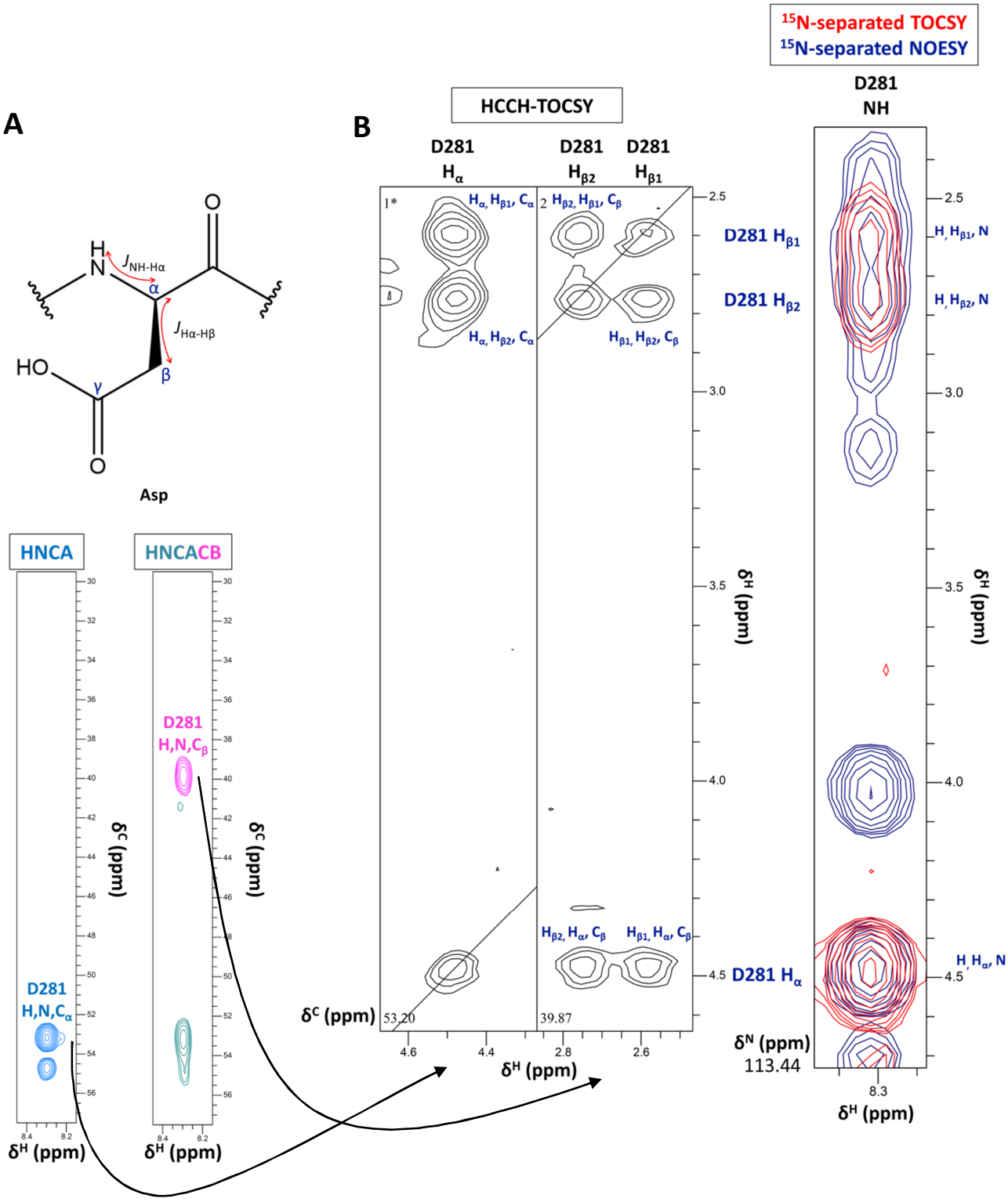


Figure 4.15 Assignment of CH groups in a short aliphatic side-chain. (**A**) Displayed formula of Asp. The key *J*-couplings are highlighted. (**B**) The Asp281 C_α and C_β peaks in the HNCA and HNCACB were used to navigate to the corresponding ^{13}C strips in the HCCH-TOCSY spectrum. The corresponding ^{15}N strip in the ^{15}N-separated TOCSY and ^{15}N-separated NOESY spectra was used in parallel to assign the side-chain proton chemical shifts. ^{15}N-separated NOESY in navy, ^{15}N-separated TOCSY in red and HCCH-TOCSY in black.

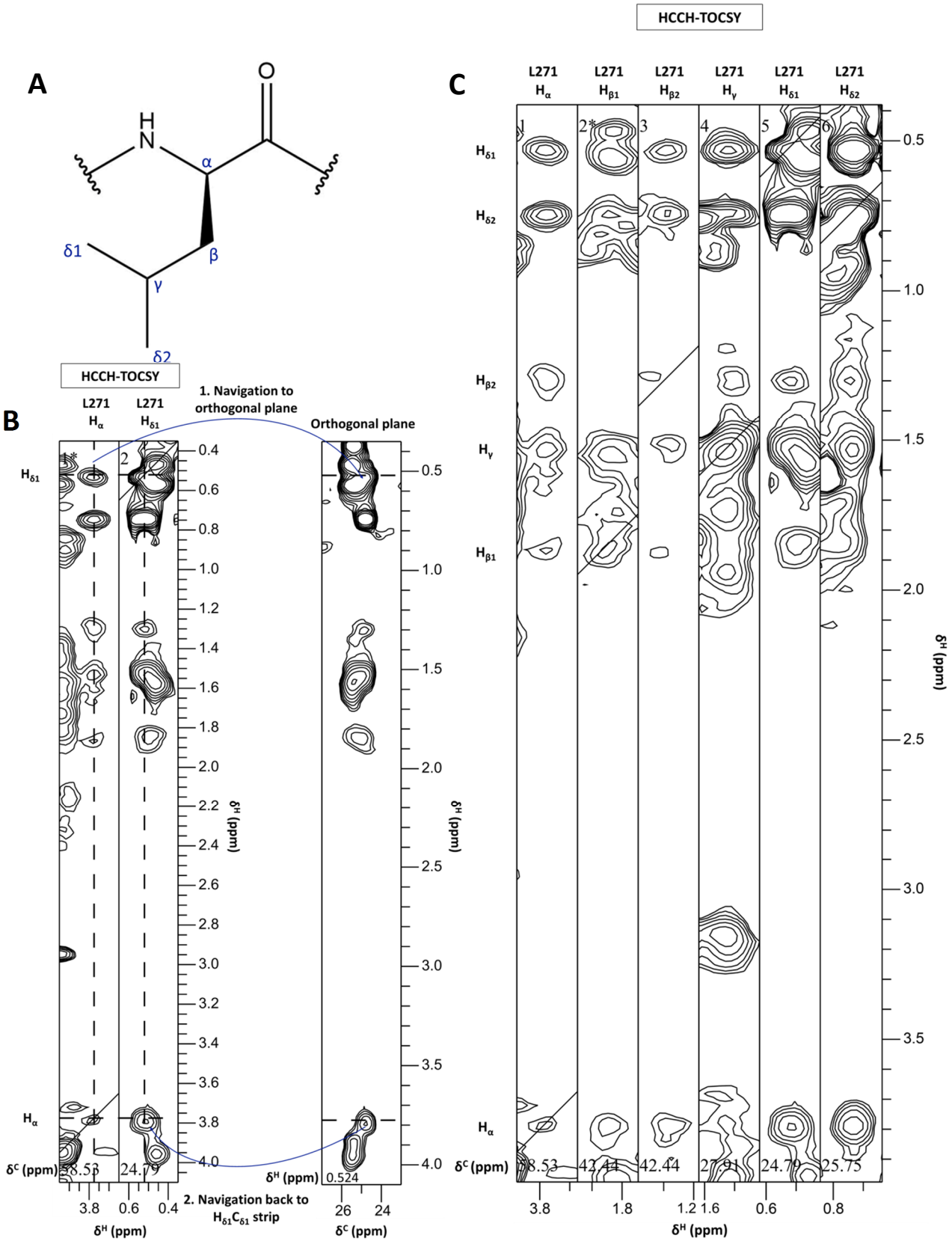


Figure 4.16 Assignment of CH groups in a longer aliphatic side-chain. (**A**) Displayed formula of Leu. (**B**) The Leu271 $H_\alpha C_\alpha$ and $H_\beta C_\beta$ peaks were assigned as shown in Figure 4.16. To assign the $C_{\delta1}$ chemical shift, the H_α $H_{\delta1}C_\alpha$ peak was used to navigate to the orthogonal plane of the HCCH-TOCSY spectrum. A matching pattern was located and then there was navigation back to the $H_{\delta1}C_{\delta1}$ strip. The same procedure was repeat for the Leu $H_\gamma C_\gamma$ and the $H_{\delta2}C_{\delta2}$. (**C**) HCCH-TOCSY strips of Leu271 $H_\alpha C_\alpha$, $H_{\beta1}/H_{\beta2}C_\beta$, $H_\gamma C_\gamma$, $H_{\delta1}C_{\delta1}$ and $H_{\delta2}C_{\delta2}$. HCCH-TOCSY in black.

4.3.8 NH_2 group assignment

There are two asparagines and three glutamines in PRK1 HR1c. As a result there are ten peaks in the ^{15}N-HSQC which correspond to the side-chain NH_2 groups. In asparagine, the NH_2 protons are close in space to the H_β protons (Figure 17A). Similarly, in glutamine, the NH_2 protons are close in space to the H_γ protons (Figure 17B). The NH_2 peaks in the ^{15}N-HSQC were used to navigate to the corresponding ^{15}N strip in the ^{15}N-separated NOESY spectrum. NOEs from the NH_2 protons to the nearest intra-residue CH group ($H_\beta C_\beta$ for Asn and $H_\gamma C_\gamma$ for Gln) protons were used to assign the NH_2 side-chain chemical shifts (Figure 4.17C). The ^{13}C-separated NOESY experiment was also used to find return NOEs from the H_α and H_β atoms to the side-chain NH_2 protons to confirm the assignments.

4.3.9 Assignment of Phe^{222}

PRK1 HR1c has a single phenylalanine residue, F222, which is in the first α-helix. The H_α and H_β chemical shifts were assigned firstly as explained above. The ^{15}N-separated NOESY and ^{13}C-separated NOESY spectra were used to help assign the H_δ, H_ε and H_ζ protons (Figure 4.18). The H_δ and H_ε are expected to have very similar chemical shifts, which makes their assignment difficult. H_δ and H_ε NOEs were assigned on the basis that the H_δ is closer to the H_β, so the NOE between these two protons should be stronger than that between the H_ε and the H_β (and also by using the NH strip where there is an NOE from the NH to the H_δ). Finally, the H_ζ was assigned by NOEs to the H_δ and H_ε.

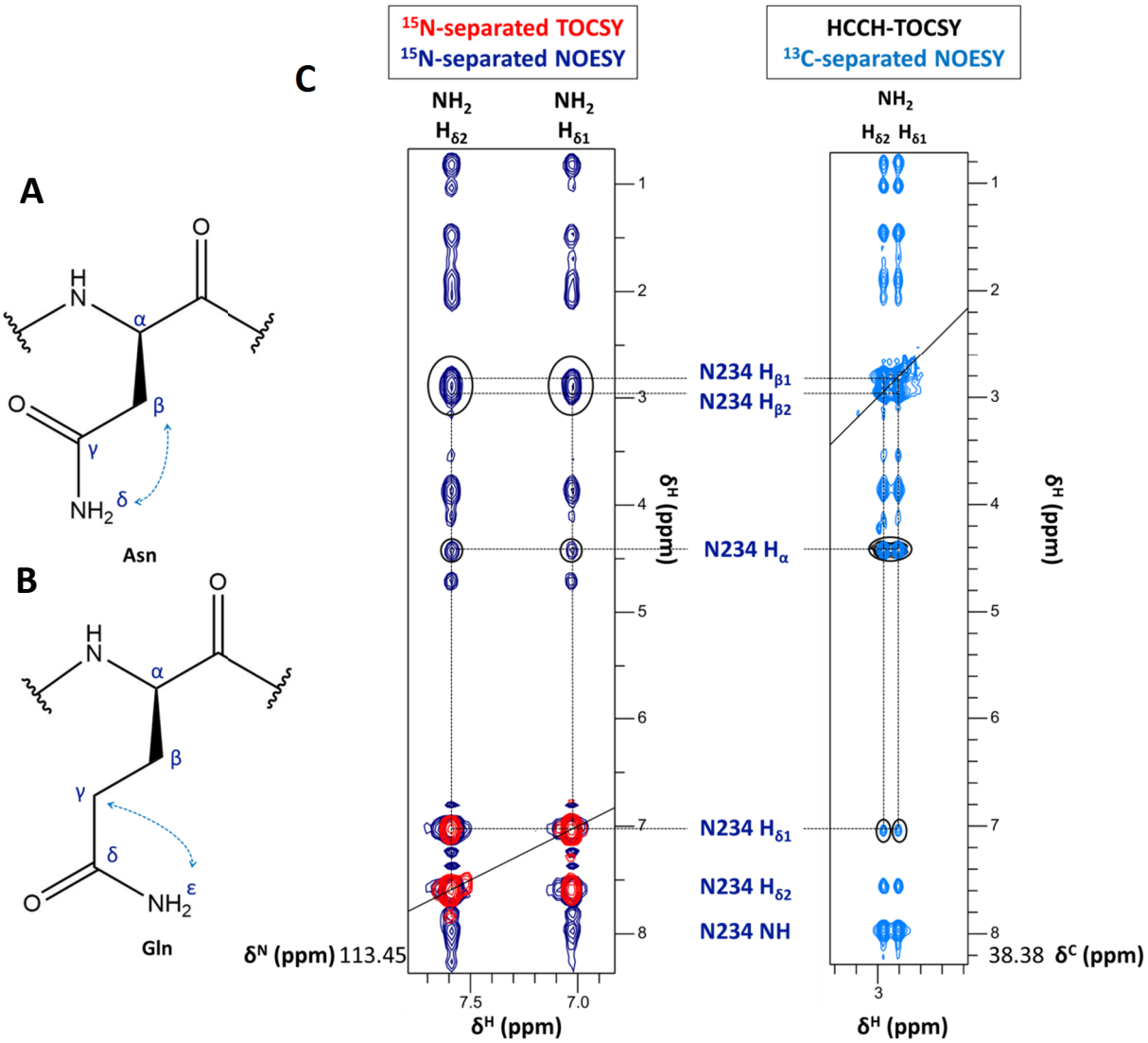


Figure 4.17 Assignment of side-chain NH_2 groups. (**A-B**) Displayed formulae of (**A**) Asn and (**B**) Gln. The arrows highlight the strongest NOE between the side-chain NH_2 protons and an intra-residue CH group. (**C**) Assignment of Asn234. NOEs to the nearest CH group protons were used to assign the side-chain NH_2 proton chemical shifts and were confirmed by return peaks in the ^{13}C-separated NOESY in the corresponding CH strip in the HCCH-TOCSY/^{13}C-separated NOESY window. These NOEs are indicated with black circles. ^{15}N-separated NOESY in navy, ^{15}N-separated TOCSY in red, HCCH-TOCSY in black and ^{13}C-separated NOESY in skyblue.

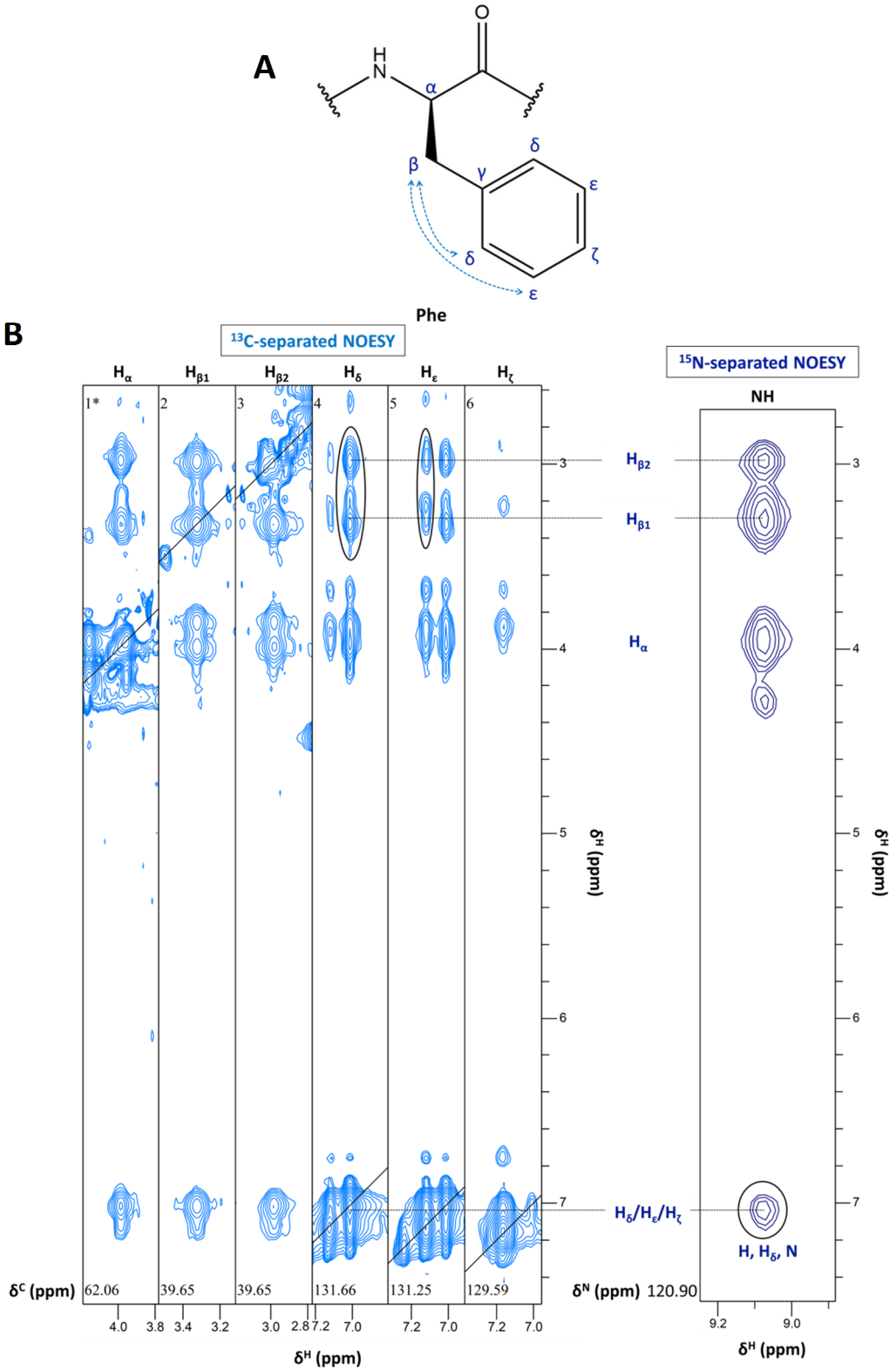


Figure 4.18 Assignment of Phe222 side-chain protons using CH-CH NOEs. (**A**) Chemical formula of phenylalanine. The arrows indicate intra-residue NOEs between the H_β and the H_δ or between the H_β and H_ε protons. (**B**) Aligned CH strips of the H_α, $H_{\beta1}$, $H_{\beta2}$, H_δ, H_ε and H_ζ protons in the ^{13}C-separated NOESY (in skyblue) and NH strip in ^{15}N-separated NOESY (navy). The strength of the H_δ and H_ε NOEs to the H_β protons was used as the basis to assign them. The black circles indicate these NOEs.

4.3.10 Assignment of prolines

The C_α and C_β chemical shifts were assigned using the triple resonance experiments as explained above. These were used to navigate to the correct C_α and C_β planes in the HCCH-TOCSY. In addition, the ^{15}N-separated NOESY was used to assign the H_δ protons on the basis of an NOE from the NH of the preceding residue to the proline H_δ protons. Furthermore, the H_δ protons of proline are near the H_α of residue i_{-1}, giving rise to strong NOE peaks between them (Figure 4.19). This was used as the basis to help assign the proline H_δ chemical shifts.

4.3.11 Medium and long range NOEs

Medium range NOEs were seen between Ha_i etc. and NH_{i+3} or side-chain protons, e.g. Thr258 H_α and Asn261 NH/H_β/H_δ (Figure 4.20).

Long range NOEs were identified by systematically going through strips of the 3D NOESY spectra and picking peaks which were identified neither as intra-residue NOEs nor as NOEs to neighbouring amino acid protons. If an already assigned proton had a very similar chemical shift, then the NOE was assigned to be to that proton (only if it was the only proton with that chemical shift). The identity of the interacting proton in these long range NOEs was confirmed by checking for a return NOE peak. This was done by selecting the NOE peak and then navigating to the orthogonal plane of the spectrum to locate the cross-peak.

4.3.12 Ambiguous and unambiguous NOEs

There is often ambiguity arising when assigning NOE peaks. This arises when an NOE peak is actually many NOE peaks which are overlapped as a result of the protons experiencing a similar chemical environment or because of poor spectral dispersion. The volume of the NOE peak is therefore dependent on all of these contributing pairs of protons. Such peaks

were often left with one or more dimensions unassigned. Particular care was taken when assigning long range NOEs, as any wrong assignments could have a deleterious effect on calculated structures. NOE assignment ambiguity was resolved during structure calculations as explained in the following section (Section 4.4).

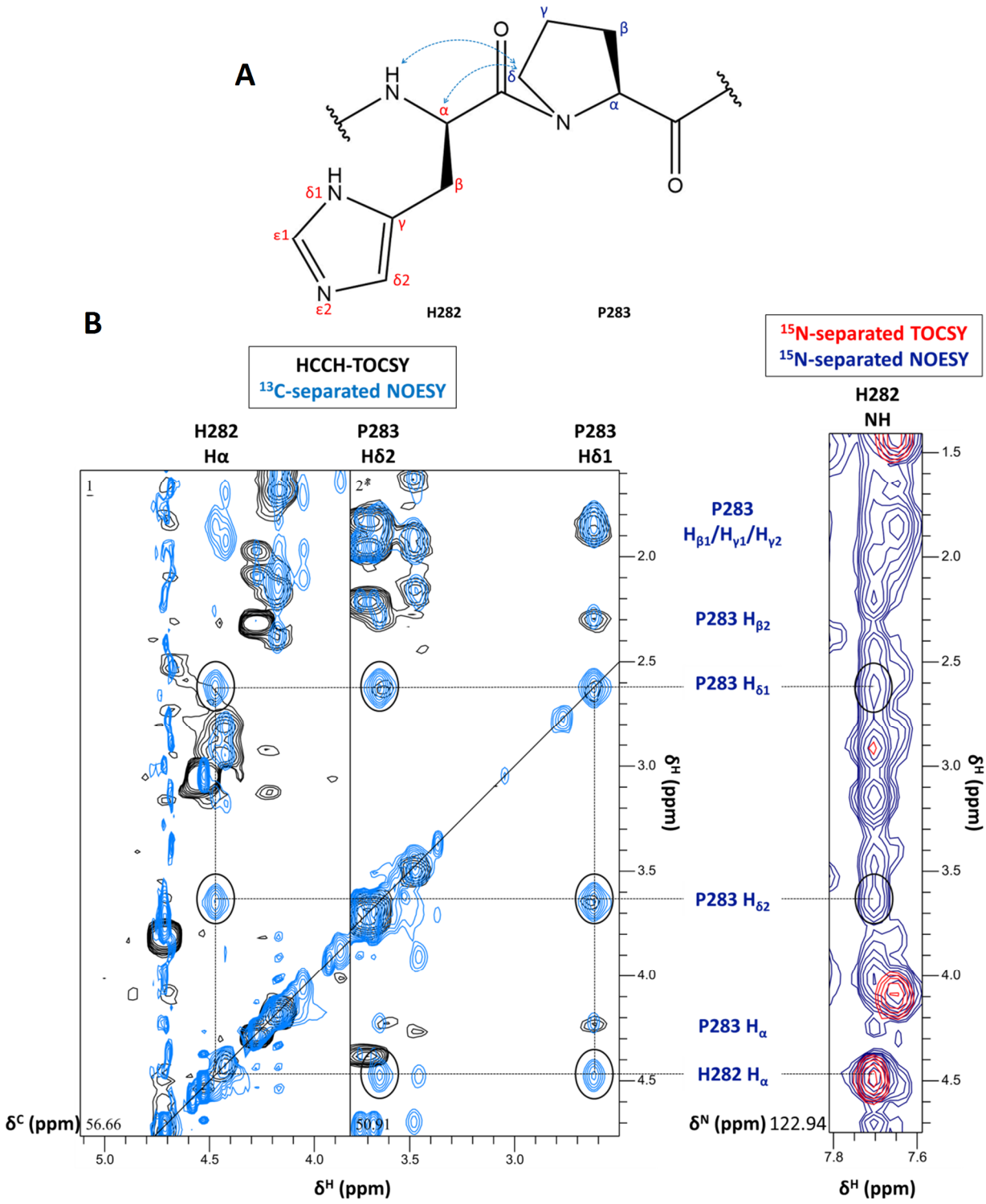


Figure 4.19 Assignment of proline side-chains. (**A**) Displayed formula of a His-Pro di-peptide. The arrows highlight the inter-residue NOEs between the proline H_δ protons and the adjacent NH and H_α protons of residue i_{-1} (His282). (**B**) The His282 H_α and Pro283 $H_{\delta1}/H_{\delta2}$ strips in the HCCH-TOCSY/^{13}C-separated NOESY spectra and the His282 NH strip in the ^{15}N-separated NOESY spectrum were used to assign the chemical shifts of proline side-chain $H_{\delta1}$ and $H_{\delta2}$ protons by identifying NOEs between them and the His282 NH and H_α protons. These NOEs are indicated with black circles. ^{15}N-separated NOESY in navy, ^{15}N-separated NOESY in red, HCCH-TOCSY in black and ^{13}C-separated NOESY in skyblue.

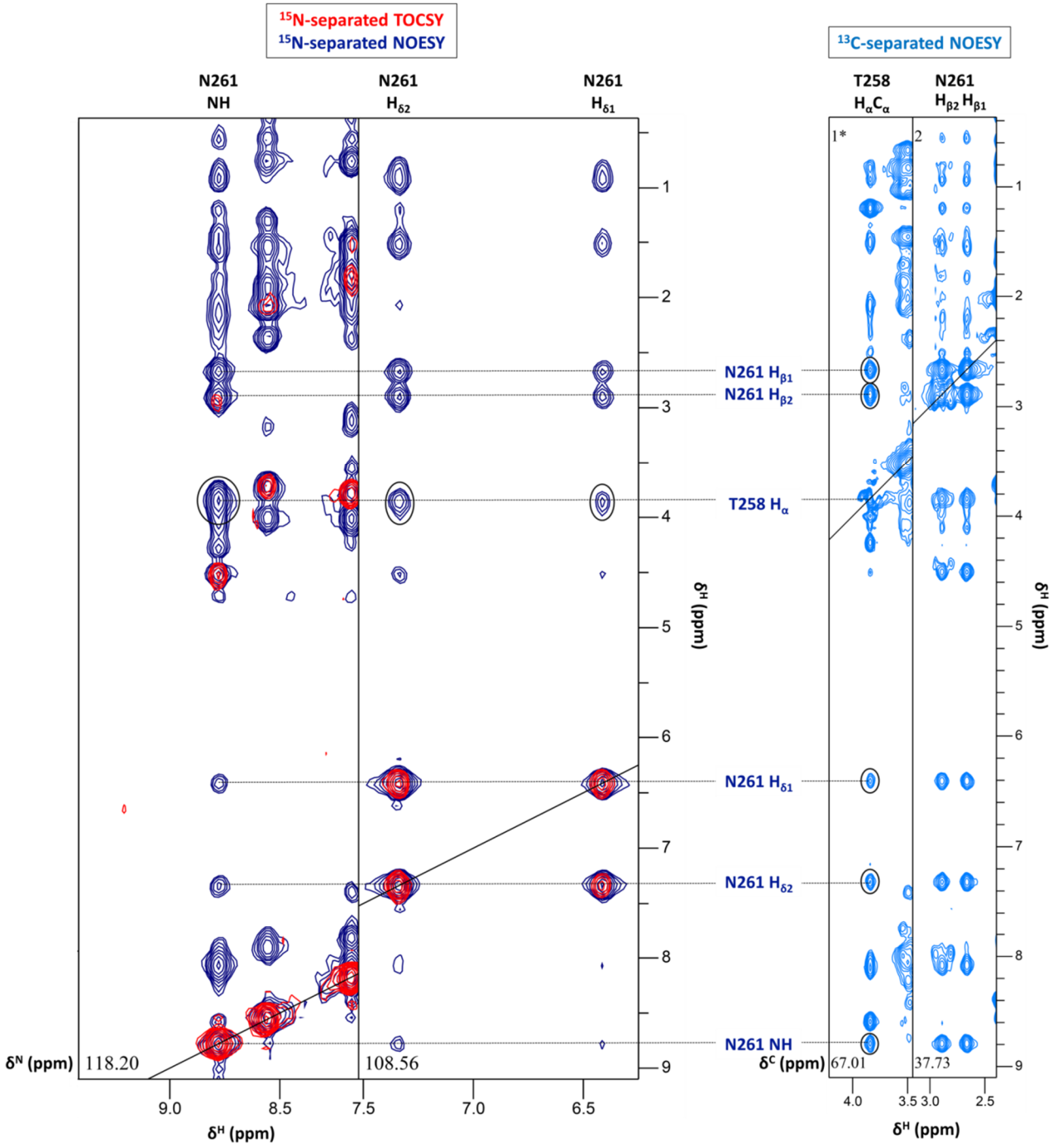


Figure 4.20 Assignment of α-helical *i* to *i+3* NOEs. The Asn261 backbone NH and side-chain NH_2 ($H_{\delta1}/H_{\delta2}$) strips in the ^{15}N-separated NOESY and the Thr258 $H_\alpha C_\alpha$ and Asn261 $H_{\beta1}/H_{\beta2}C_\beta$ strips in the ^{13}C-separated NOESY were used to assign α-helical i to i+3 NOEs. These NOEs are indicated with black circles. ^{15}N-separated NOESY in navy, ^{15}N-separated TOCSY in red and ^{13}C-separated NOESY in skyblue.

4.4 Structure calculations using ARIA

4.4.1 Dihedral angles

A key restraint in structural calculations is the set of dihedral or torsion angles, ϕ and ψ of the protein polypeptide backbone. While NMR methods are available to measure the dihedral angles of peptides, this is not feasible when work is done on larger proteins. TALOS (Torsion Angle Likeliness Obtained from Shift and Sequence Similarity), and its newer versions named TALOS+ and TALOS-N can predict protein local geometry from chemical shifts (Shen & Bax, 2013), as chemical shifts are inherently related to the local protein structure.

TALOS-N takes the input chemical shifts and creates dihedral angle restraint lists by comparing to two protein structure databases. The first database comprises 580 proteins for which both backbone chemical shifts (^{15}N, ^{13}C, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{1}H^{\alpha}$ and $^{1}H^{N}$) are available in the BMRB (Markley *et al.*, 2008) and atomic coordinates are available from high-resolution X-ray structures in the PDB (Berman *et al.*, 2012). The second database utilised by TALOS-N contains 9523 high-resolution X-ray structures for which chemical shift assignments are not experimentally available but rather predicted using the SPARTA+ program (Shen & Bax, 2010). SPARTA+ performs the reciprocal of TALOS-N, i.e. determines chemical shifts from local geometry. Specifically, a ϕ/ψ distribution code is assigned to every residue of each protein in this database, linking the backbone conformation of each residue to the Ramachandran plot. An artificial neural network then links the ϕ/ψ distribution code to chemical shift. This highlights why the database uses a large number of high-resolution structures. Put together, the two databases utilised by TALOS-N establish a link between chemical shift and local geometry.

TALOS-N utilises an artificial neural network (based on the first database) to predict ϕ/ψ distributions for each residue of the protein in question from submitted NMR chemical shifts, the residue type and adjacent residues. The protein is then split into heptapeptides and the second, larger database is used as a reference to identify heptapeptides with matching ϕ/ψ distributions and preferably heptapeptides with a matching central residue.

The top 1000 heptapeptides are chosen. The SPARTA+ predicted chemical shifts of these 1000 peptides are then scored and further compared with the query protein experimental chemical shifts. The 25 top matches are found and the ϕ/ψ distribution of their middle residue is averaged. This ϕ/ψ distribution average is used as the prediction of the ϕ/ψ distribution of the protein heptapeptide.

In brief, NMR chemical shifts are submitted to TALOS-N which goes on to predict ϕ/ψ distributions for every residue based on the first, smaller database of protein structures. The protein is then divided into heptapeptides and the top ϕ/ψ distribution matches are selected from the second, larger database of protein structures. The matching heptapeptides are further refined on the basis of their predicted chemical shifts (by SPARTA+) with the experimental chemical shifts of the query protein. The chemical shifts of these top 25 matches are then used to calculate an average ϕ/ψ distribution which is used as the predicted ϕ/ψ distribution for the query protein. The predicted ϕ/ψ distribution combines data from both databases and is therefore less prone to error if only the smaller database were used to predict ϕ/ψ angles directly. Additionally, TALOS-N is an improvement over the original TALOS and TALOS+ as it uses heptapeptides (and not tripeptides) which are more representative of local protein structure. Predicted dihedral angles for each residue provided by TALOS-N can therefore be used as restraints in structure calculations.

In addition to dihedral angle predictions, TALOS-N also classifies these predictions. Specifically, if the ϕ/ψ angles of the central residues of each of the 25 top matched heptapeptides cluster in the same region of the Ramachandran plot, then the prediction is classified as “strong”. In contrast, if more than 10 but less than 25 ϕ/ψ angle values cluster in the same region, the prediction is classified as “generous”, while if there is no clear clustering the prediction is classified as “warning”, i.e. the prediction is ambiguous. Residues may also be classified as “dynamic” if they are predicted to have a dynamic conformation.

PRK1 HR1c chemical shifts (see Appendix 4) were submitted to TALOS-N to obtain dihedral angle predictions. Residues 208-241 and 243-294 were classified as strong predictions and 207 as a generous prediction. 195-197, 202-206 and 295-296 were classified as dynamic while 198-201 and 242 were classified as ambiguous.

4.4.2 Distance restraints

The other important structure calculation restraints which come from our data are distance restraints. These are based on the strength of NOE peaks selected in NOESY spectra. The greater the volume (or intensity) of the peak, the closer two protons are in space. The NOE is useful for any protons which are up to 5 Å away. As explained in section 4.3.12, there is often peak overlap in NOESY spectra and sometimes the identity of all the protons contributing to a particular NOE peak is unknown. ARIA (Ambiguous Restraints for Iterative Assignment) is a structure calculation method which gets around this problem and facilitates structure calculations. The ambiguity in the NOE assignments is resolved during the structure calculation. The method will be discussed here briefly.

As stated above, the volume of an NOE peak between two spins, V_{ij}, is proportional to the distance between them, d_{ij}

$$V_{ij} \propto {d_{ij}}^{-6} \qquad 4.20$$

but this relationship is also subject to protein concentration, dynamics and spin diffusion (Nilges, 1996). As a result, the constant of proportionality relating distance to the NOE is unknown. ARIA gets around this problem by performing distance calibration during the structure calculations (Linge *et al.*, 2001) and for this reason this constant can be written as the calibration factor, C_{cal}. In the first iteration of the structure calculations, iteration zero, this factor is derived from a structure which has extended geometry with randomised torsion angles and side-chains with idealised geometries. As a result, C_{cal} is subject only to distances between intra-residue spins. With subsequent iterations, C_{cal} is derived from the lowest energy calculated structures and refined after every single iteration. Since the distance is calibrated with each subsequent iteration, we can write

$$d^{obs} = (C_{cal} V_{ij})^{-\frac{1}{6}} \qquad 4.21$$

where d_{obs} is the calibrated distance.

ARIA gets around ambiguous assignments by describing the effective distance, D, which describes the summation of distances (d_i) between the proton pairs of all possible assignments which total N_δ, where δ is the chemical shift tolerance

$$D \equiv \left(\sum_{i=1}^{N_\delta} {d_i}^{-6} \right)^{-\frac{1}{6}} \qquad 4.22$$

We can then state that the ambiguous NOE strength will be proportional to the inverse sixth power of the effective distance

$$V_{ij} \propto D^{-6} \equiv \left(\sum_{i=1}^{N_\delta} {d_i}^{-6} \right) \qquad 4.23$$

The ARIA method uses 8 consecutive iterations in the structure calculation. With subsequent iterations there is an increasing number of ambiguous NOEs which are assigned. The final iteration involves refinement of the structure in water.

4.4.3 Initial structure calculations

A total of 2500 distance restraints were submitted to ARIA for the first structure calculation, with approximately 50% of these being unambiguous. The structure calculation ran for a total of 8 iterations, with 20 structures calculated during each iteration. The 10 lowest energy structures were refined in water and used to produce a list of restraint violations.

4.4.4 Manual refinement

Restraint violations were manually inspected and key problems, e.g. swapped assignments of Arg247 and Arg274, were obvious and corrected immediately. Furthermore, many violations were the product of very weak NOEs, e.g. due to spin diffusion, being used in the structure calculations. Removal of these weak NOEs, picking and assigning more NOEs and

manual refinement of the key violations were then performed and the next structure calculation was set up.

With each new structure, the number of violations decreased, e.g. from 600 to 500 and so on. In addition, as the key violations were sorted out, other smaller violations also disappeared.

Structures calculated initially suffered from wrongly assigned ambiguous NOEs. This caused helix 1 of PRK1 HR1c bending from its N-terminus upwards. With each subsequent structure calculation, the ensemble of structures improved and could then also be used to derive distance information to unambiguously assign more NOEs.

A total of 100 structures were calculated in the final iteration of the last structure calculation. The 50 lowest energy structures were refined in water and the structure ensemble consisted of the 45 lowest energy structures (section 4.4.5).

The average structure was calculated from the structure ensemble and the rmsd of each of the 45 structures from the average was then determined. This was used to identify the structure with the lowest rmsd, i.e. the one that is closest to the average structure. Table 4.1 includes the key structure statistics of both the ensemble of structures and the structure closest to the mean. For the structure ensemble, the backbone rmsd of the ordered regions was 0.552 while the sidechain rmsd was 0.989. The statistics also include information on dihedral angles based on a Ramachandran plot generated by the program PROCHECK.

Experimental restraints		
Completeness of chemical shift assignment (%):		
C	70.8	
H	98.4	
N	65.8	
NOE-based distance restraints:		
Total non-degenerate	2085	
Unambiguous	1534	
Ambiguous	551	
Violations		
NOE distance violations ≥ 0.5 Å	1.9 ± 1.59	
Torsion angles ≥ 5 °	0.66 ± 0.710	
Average Van der Waal energy (kcal/mol)	-177 ± 16.9	
	A	**B**
Coordinate precision for well-ordered regions		
RMSD of backbone atoms (Å)	0.552 ± 0.174	0.255
RMSD of all heavy atoms (Å)	0.989 ± 0.146	0.793
Ramachandran analysis for all residues		
Residues in most favoured regions (%):	92.9	95.6
Residues in additionally allowed regions (%):	6.3	3.3
Residues in generously allowed regions (%):	0.5	0
Residues in disallowed regions (%):	0.3	1.1
Ramachandran analysis for ordered regions		
Residues in most favoured regions (%):	98.4	95.6
Residues in additionally allowed regions (%):	1.5	3.3
Residues in generously allowed regions	0.1	0
Residues in disallowed regions (%):	0	1.1
Sidechain analysis for ordered regions		
Rotomeric sidechains (%):	97.1	97.1
Non-rotomeric sidechains (%):	2.9	2.9
RMS deviations for all residues		
from experimental restraints:		
NOE distances (Å)	0.0524 ± 0.0148	0.0351
Dihedral angles (°)	0.875 ± 0.105	0.911
from idealised geometry:		
Bond lengths (Å)	0.00446 ± 0.000269	0.00410
Angles (°)	0.615 ± 0.0223	0.580
Impropers (°)	1.38 ± 0.128	1.23

Column A contains information on the ensemble of 45 structures.
Column B contains information for the structure that is closest to the mean.

Table 4.1 Experimental restraints and structural statistics.

4.4.5 The structure

The PRK1 HR1c domain consists of three α-helices (Figure 4.21). The first two helices form an antiparallel coiled coil structure which resembles the structure of other HR1 domains. Helix 1 spans 34 residues (Leu208-Ala241) while helix 2 spans 31 residues (Arg247-Glu277). Helix 3 is a C-terminal helix which spans 12 residues (Lys284-Ala295) in a total of 3.6 α-helical turns. Two inter-helical loops exist, connecting helices 1 and 2 and helices 2 and 3, respectively. The first inter-helical loop comprises Ala242-Asp246 while the second inter-helical loop comprises Leu278-Pro283. The N-terminal region of the protein is unstructured and this is in agreement with the TALOS-N secondary structure prediction. Overall, the secondary structure in the structure matched the prediction from TALOS-N.

Helix 2 is fairly straight while helix 1 twists around it (slight left-handed twist). Helix 3 forms contacts with both the N-terminal part of helix 1 and the C-terminal part of helix 2 to form a triple helix, while it slightly twists around the other two helices.

The structure is dependent on short (intra-residue), medium (i to i+3) and long range NOEs. The antiparallel coiled coil between helices 1 and 2 is formed by multiple, hydrophobic inter-helical contacts between side-chains of Ile215, Leu218, Phe222, His226, Ala229, Ala232, Leu236, Leu239 (helix 1) and side-chains of Val250, Ala253, Leu257, Asn261, Leu264, Leu267, Leu271 and Leu275 (Figure 4.22). At the same time, helix 3 is also involved in hydrophobic interactions with helices 1 and 2. For example, the H_δ of Leu288 (helix 3) makes a Van der Waals contact with the H_δ of Leu275 (helix 2).

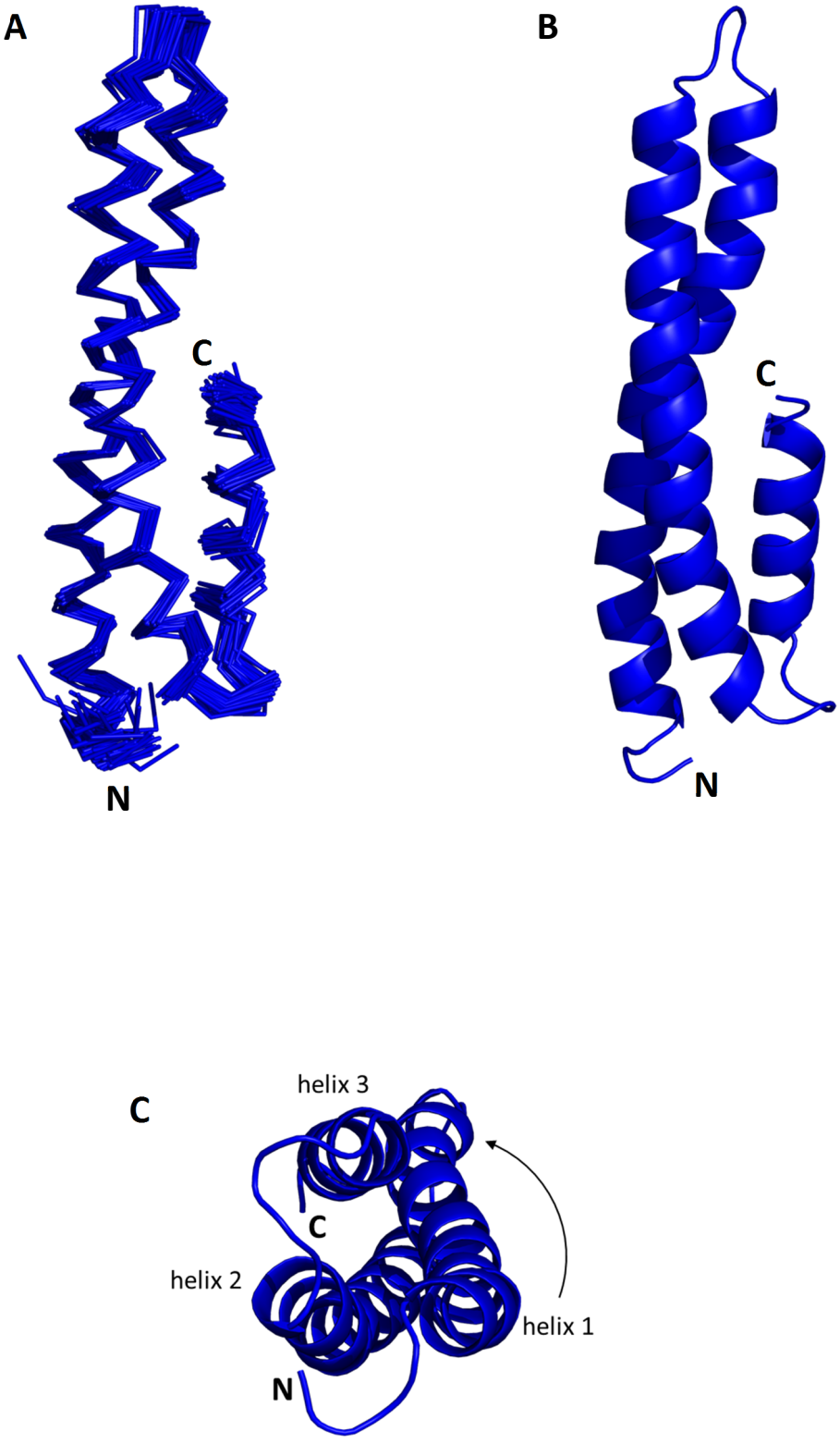


Figure 4.21 The NMR solution structure of PRK1 HR1c. (**A**) The ensemble of the 45 lowest energy structures of PRK1 HR1c. The backbone traces are overlaid. (**B**) Cartoon representation of the lowest energy structure. (**C**) End-on view of the protein in cartoon representation. Helix 1 wraps around helix 2 with left-handed twist. Figures were made with PyMol.

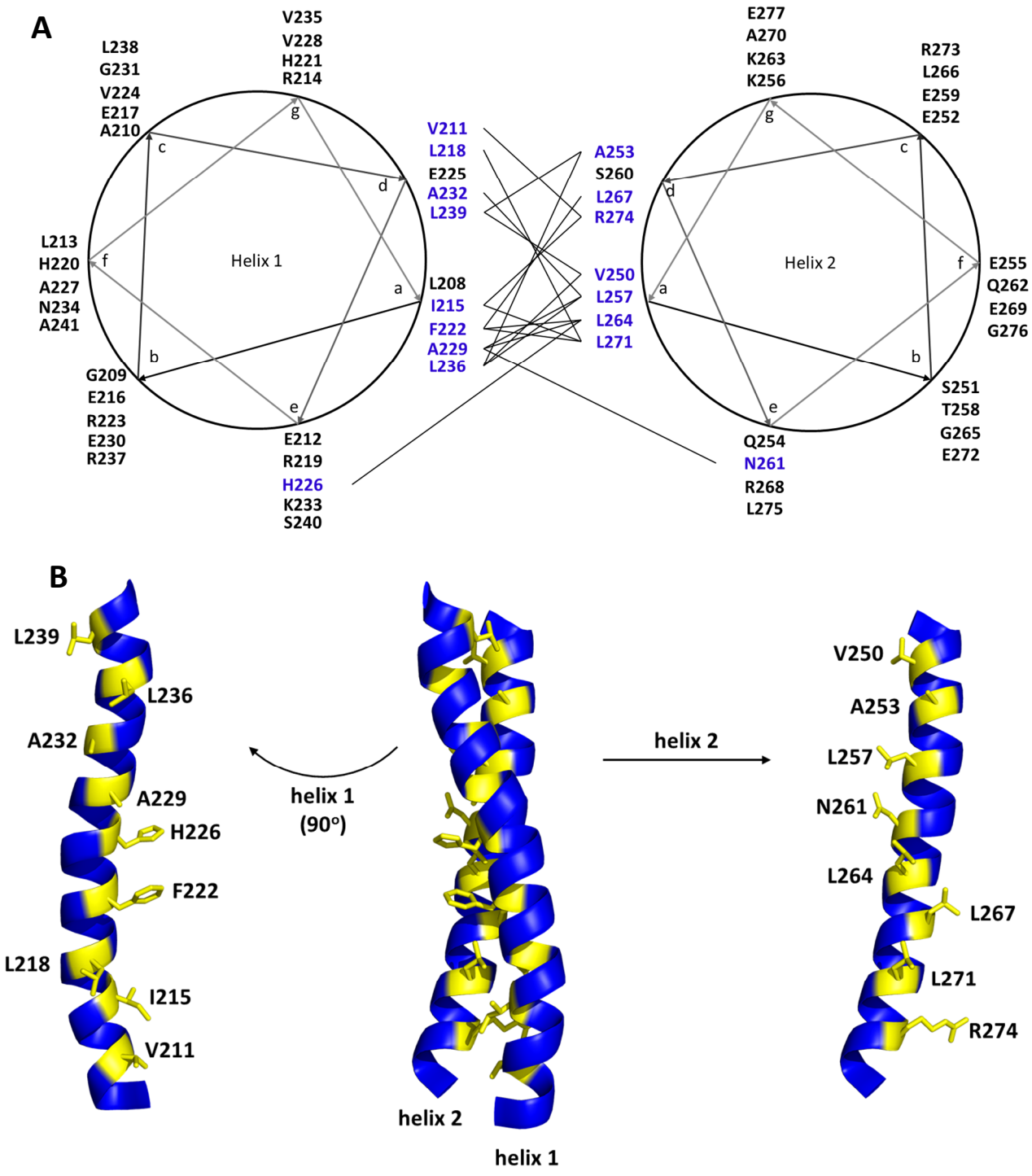


Figure 4.22 Inter-helical contacts in the coiled coil. (**A**) Helical wheel plot of the PRK1 HR1c coiled coil. Residues involved in inter-helical contacts are highlighted in blue. (**B**) Cartoon representation of the PRK1 HR1c coiled coil (in blue). Residues involved inter-helical contacts are highlighted in yellow and the side-chains are shown as sticks. The inter-helical contacts are also shown on each individual helix, highlighting the twist in the coiled coil.

The triple coiled coil region consists of interactions between all three helices. Out of the three helices, helix 2 contributes more side-chains to these interactions than helices 1 or 3 do. The most extensive interactions involve the phenyl ring of Phe222 (helix 1) (Figure 4.23), which is buried between the methyl groups of Leu264, Leu267 and Leu271 (helix 2). Importantly the overall twist of helix 1 appears to favour docking of the phenyl ring in the hydrophobic pocket made by these three helix 2 leucines. The phenyl ring also makes hydrophobic contacts with the side-chain methyl group of Leu292 (helix 3) and with the H_{β} of Ala295 (helix 3), promoting the positioning of helix 3 alongside helices 1 and 2. In addition, hydrophobic contacts are formed between Ile215 (helix 1), Leu271 (helix 2) and Leu288 (helix 3) which further promotes formation of the triple coiled coil structure.

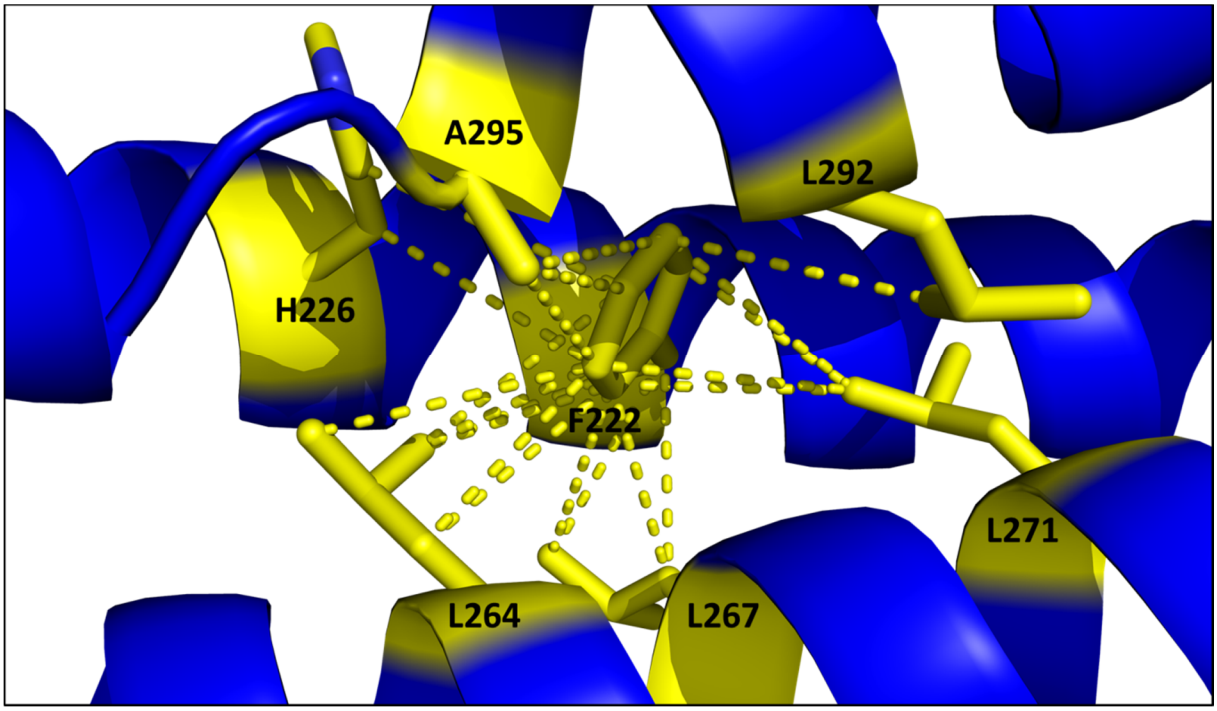


Figure 4.23 The key long-range NOEs in the triple helical region. A cartoon representation of the triple helical region of PRK1 HR1c in blue. Residues involved in NOEs to Phe222 are shown as sticks and are coloured in yellow. Dotted yellow lines indicate the key NOEs. Phe222 (helix 1) is buried in a hydrophobic pocket made by Leu264, Leu267 and Leu271 (helix 2). Leu292 and Ala295 (helix 3) and His226 (helix 1) also make contacts with Phe222.

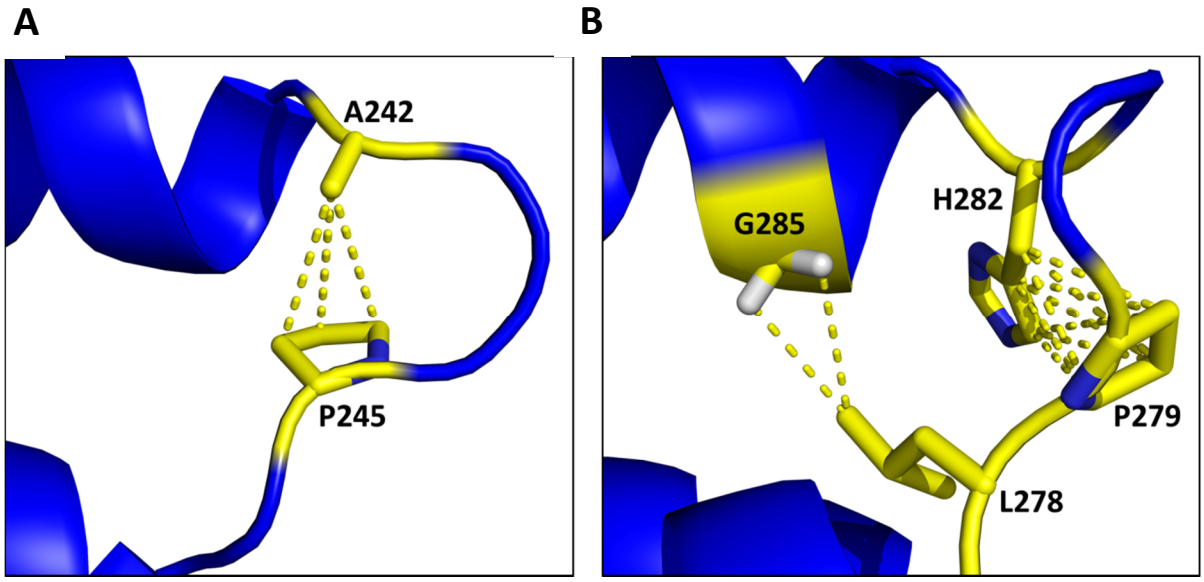


Figure 4.24 The loop region NOEs. A cartoon representation of the inter-helical loops in PRK1 HR1c. PRK1 HR1c is in blue while residues with key NOEs are coloured in yellow. Dotted yellow lines indicate the key NOEs. (**A**) Loop between helices 1 and 2. The side-chain methyl group of Ala242 contacts the side-chain of Pro245. (**B**) Loop between helices 2 and 3. The H_{δ} of Leu278 contacts the H_{α} of Gly285 (helix 3), while multiple NOEs exist between the Pro279 and His282 side-chains.

Important NOEs were also observed in the two inter-helical loops (Figure 4.24). The turn in the first inter-helical loop is mediated by the contact between the side-chain methyl group of Ala242 and the Pro245 side-chain. The remaining side-chains are solvent exposed. In the second inter-helical loop, the side-chain methyl of Leu278 (just after the end of helix 2) makes a Van der Waals contact with the H_{α} Gly285 (helix 3), helping to position helix 3 against helix 2. The turn made by this inter-helical loop is dependent on contacts between the side-chains of Pro279 and His282.

4.4.6 Comparison with other HR1 domain structures

Structures of other HR1 domains have been solved and the HR1 domains all form anti-parallel coiled coils. These HR1 domains are PRK1 HR1a (Maesaki *et al.*, 1999; PDB:1CXZ), PRK1 HR1b (Owen *et al.*, 2003; PDB:1URF), TOCA1 HR1 (Watson *et al.*, 2016; PDB:5FRG) and CIP4 HR1 (Kobashigawa *et al.*, 2009; PDB:2KE4). The structure of PRK1 HR1a is from a crystal structure in complex with RhoA (the GTPase may have an effect on the structure of the HR1a domain), while all other HR1 domain structures are NMR solution structures just like that of PRK1 HR1c presented here.

All HR1 domains have two main α-helices which form an antiparallel coiled coil (Figure 4.25). The length of these helices however varies between the different HR1 domains. The length of helix 1 (38 residues in HR1a, TOCA1 and CIP4, 30 residues in HR1b and 34 residues in HR1c) is less variable than the length of helix 2 (25, 32, 31, 29 and 35 residues for HR1a, HR1b, HR1c, TOCA1 and CIP4, respectively). For HR1a, HR1c, TOCA1 and CIP4, helix 1 is longer than helix 2, with the difference in the number of residues between the two helices ranging from 3 to 13. In contrast, helix 2 of HR1b is longer than helix 1, although only by 2 residues. The difference in length of helices 1 and 2 is greatest in HR1a and TOCA1 while it is much smaller in the case of HR1b, HR1c and CIP4. The length of the inter-helical loop also differs between the HR1 domains. While the inter-helical loop of all three PRK1 HR1 domains comprises 5 residues, the loop of TOCA1 and CIP4 comprises of 10 and 9 residues, respectively. In addition, while the inter-helical loop of the PRK1 HR1 domains is unstructured, short 3_{10} helices form in the loop of TOCA1 and CIP4. These subtle differences in structure, including the length and twist of helices as well as differences in inter-helical loops may explain the differential affinity of these HR1 domains for Rho GTPases.

In all 5 HR1 domains, helix 2 is fairly straight while helix 1 twists (left handed-twist) around helix 2 to form the coiled coil structure. As the helices have variable length and are aligned at different angles with respect to each other, the resulting twist is not the same in the 5 HR1 domains. This results in an inter-helical distance of approximately 10 Å at both coiled coil ends of HR1c, TOCA1 and CIP4. Contrastingly, in HR1a and HR1b the distance ranges from 10 to 12 Å.

HR1a and HR1c are special in that they have additional α-helices. HR1a has an N-terminal helix which comprises 5 residues and forms hydrophobic contacts with both of the main helices. The extra helix in HR1c is C-terminal and not N-terminal, but like the N-terminal HR1a helix, it also packs against the main two helices. A key difference between these two additional helices is their orientation. The C-terminal HR1c helix is slightly tilted with respect to the main two helices, while the N-terminal HR1a helix is much more tilted. The N-terminal helix of HR1a has not been shown to actually exist in solution and may be an artefact of crystal packing. In contrast, the C-terminal helix of HR1c exists in solution. It is possible that it has a role as a stabilising helix in the interaction of the other two helices with a GTPase, similar to the stabilising helix of mDia1 which helps the protein's two main helices interact with RhoC (Rose *et al.*, 2005). Whether this helix is truly fixed to the other two helices in solution or whether it can freely move around was investigated in the next section by recording and analysing relaxation experiments.

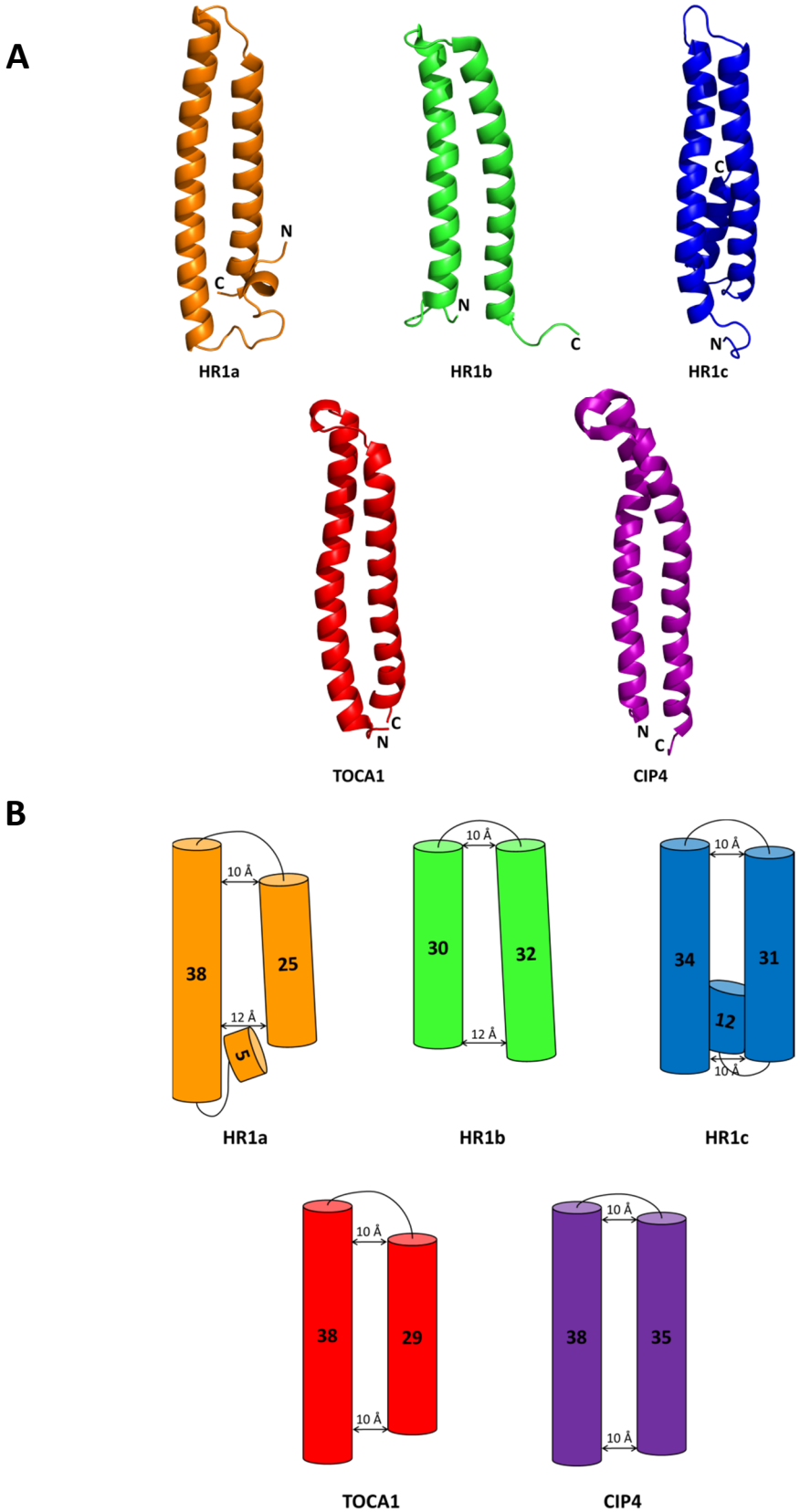


Figure 4.25 Structural comparison of HR1 domains. PRK1 HR1a (PDB:1CXZ) is in orange, PRK1 HR1b (PDB:1URF) is in green, PRK1 HR1c is in blue, TOCA1 HR1 (PDB:5FRG) is in red and CIP4 HR1 (PDB:2KE4) is in purple. (**A**) Structures of HR1 domains in cartoon representation. (**B**) The helices of each structure are represented by cylinders. The number of residues of each helix and the inter-helical distance at either end of the coiled coil are noted.

4.5 Relaxation analysis of the HR1c structure

Relaxation is the process by which the bulk magnetisation of spins returns to equilibrium, i.e. it re-aligns with the z-axis. The rate of relaxation is dependent on the environment of the relevant nuclei as well as the motion of the entire molecule. Therefore, relaxation analysis is invaluable in providing such information. ^{15}N relaxation experiments (T_1, T_2 and heteronuclear steady-state NOE experiments) are useful as they can provide information on ps-ns time-scale dynamics of backbone amides. To understand what these experiments measure it is imperative to discuss the theory behind them.

4.5.1 Theory

4.5.1.1 Transverse and longitudinal relaxation

Transverse relaxation is the process by which magnetisation in the xy-plane decays away to zero. Conversely, longitudinal relaxation is the re-appearance of the z-magnetisation to its equilibrium value. The rates of longitudinal and transverse relaxation, R_1 (or R_z) and R_2 (or R_{xy}), respectively, are defined as

$$R_1 = \frac{1}{T_1} \qquad 4.24$$

$$R_2 = \frac{1}{T_2} \qquad 4.25$$

where T_1 and T_2 are the longitudinal and transverse time constants, respectively.

4.5.1.2 Relaxation mechanisms

Two main relaxation mechanisms exist for spin-half nuclei, chemical shift anisotropy and the dipolar mechanism, and they will both be briefly discussed here.

As discussed earlier in this Chapter, the magnetic moment of each nucleus interacts with the applied magnetic field, B_0, to give rise to a small magnetic field which in turn defines the

chemical shift, δ. The size of this small magnetic field and consequently the size of δ are dependent on each molecule's orientation with respect to B_0 (and also depends on the γ of each nucleus), i.e. δ is anisotropic. As the molecules tumble in solution, there is variation in their small magnetic field and it is this variation which can lead to relaxation.

The second relaxation mechanism is the dipolar one, also known as dipole-dipole relaxation, which describes how spins can often experience the magnetic moment of a nearby spin. The local field experienced due to this nearby spin is inversely proportional to the cubic distance between the two spins (i.e. only occurs for neighbouring spins, typically <5 Å). The field is also dependent on the γ of both local field-generating and local field-experiencing spins (higher for ^{1}H), as well as the orientation of the vector connecting the two spins with respect to B_0.

4.5.1.3 The correlation and the spectral density functions and T_1 and T_2 experiments

In an NMR sample, the local magnetic field experienced by a spin, B_{local}, is not constant but rather it randomly varies with time. The correlation function, $G(t,\tau)$ describes how B_{local} varies between time t and time $t+\tau$,

$$G(t,\tau) = \overline{B_{local}(t) \times B_{local}(t+\tau)} \qquad 4.26$$

with the vinculum indicating that this is strictly representative of a population ensemble of spins.

The molecule's rotational diffusion is described by the correlation function

$$G(\tau) = \overline{B_{local}^2} \times e^{\left(-\frac{|\tau|}{\tau_c}\right)} \qquad 4.27$$

where τ_c is the rotational correlation time (the average time it takes for a molecule to rotate by one radian). Figure 4.26 shows plots of $G(\tau)$ at different τ_c values.

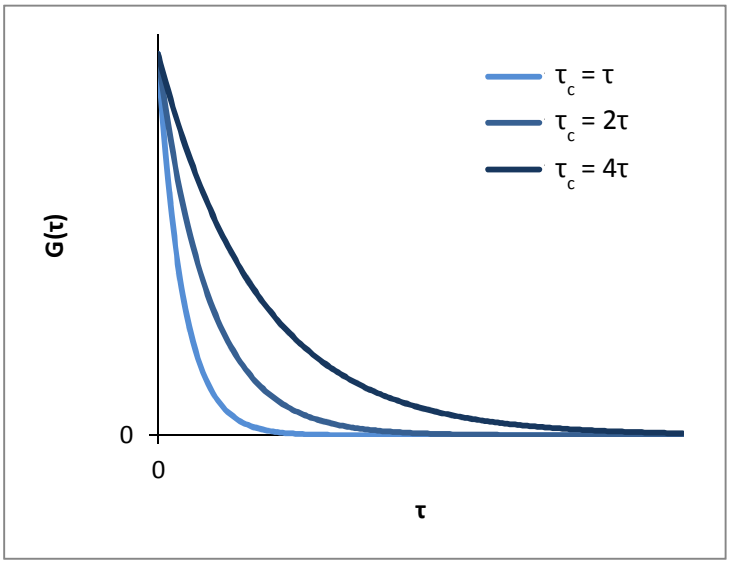


Figure 4.26 Correlation function plot. Plots of the correlation function at doubling values of the correlation time, τ_c. In all three plots, the maximum value is at $\tau = 0$, $G(t) = \overline{B_{local}^2}$.

For changes in local magnetic fields to cause relaxation, they must be oscillating near the Larmor frequency. It is therefore useful to Fourier transform the correlation function to obtain a frequency function, which is known as the spectral density function, $J(\omega)$

$$G(\tau) = \overline{B_{local}^2} \times e^{\left(-\frac{|\tau|}{\tau_c}\right)} \overset{FT}{\rightarrow} \overline{B_{local}^2} \times \frac{2\tau_c}{1+\omega^2\tau_c^2} \qquad 4.28$$

$$J(\omega) = \overline{B_{local}^2} \frac{2\tau_c}{1+\omega^2\tau_c^2} \qquad 4.29$$

The Fourier transform changes the exponential decay function $G(\tau)$ into the Lorentzian lineshape $J(\omega)$. Figure 4.27 shows a plot of the spectral density function at three different values of τ_c.

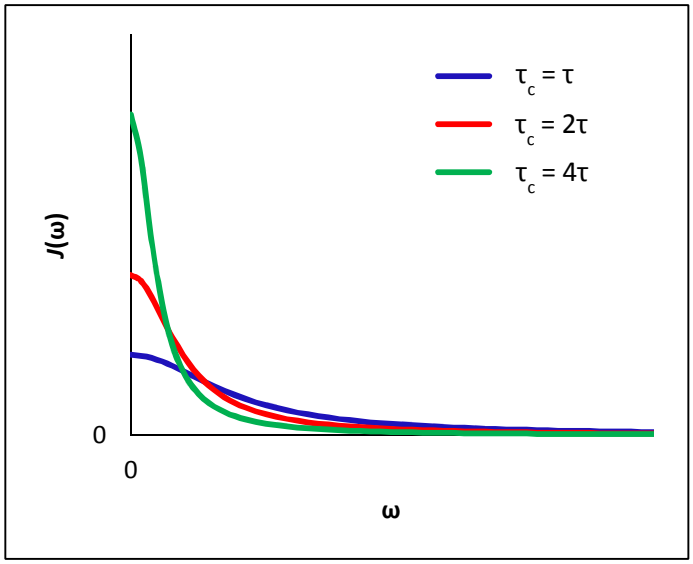


Figure 4.27 Spectral density function plots. Plots of $J(\omega)$ at doubling values of the correlation time, τ_c. $2\overline{B_{local}^2}\tau_c$ is the maximum (at $\omega = 0$).

The higher the value of τ_c, the lower the frequencies to which $J(\omega)$ effectively spreads. An important point to note is that by integration the area under $J(\omega)$ is $\pi\overline{B_{local}^2}$, i.e. it is independent of τ_c. This practically means that with decreasing τ_c, the maximum value of $J(\omega)$ (at $\omega = 0;\ \omega_0$), also decreases. This is important as the rate of longitudinal relaxation is dependent on the spectral density at the Larmor frequency, $J(\omega_0)$, which can be written as

$$J(\omega_0) = \overline{B_{local}^2}\frac{2\tau_c}{1+\omega_0^2\tau_c^2} \qquad 4.30$$

and is shown as a plot of $J(\omega_0)$ against τ_c in Figure 4.28. The maximum of this function is given when $\tau_c = \frac{1}{\omega_0}$, and it is at this value that longitudinal relaxation is the greatest.

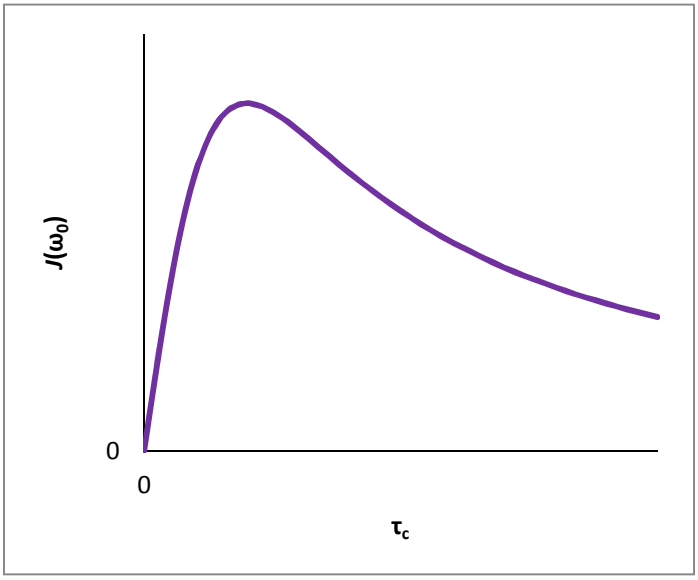


Figure 4.28 Spectral density function at Larmor frequency. Plot of $J(\omega_0)$ against the rotational correlation time, τ_c. The maximum is given when $\tau_c = \frac{1}{\omega_0}$.

The reduced spectral density $j(\omega)$ is the simplified version of $J(\omega)$ which is independent of local fields and is written as

$$j(\omega) = \frac{2\tau_c}{1 + \omega^2\tau_c^2} \qquad 4.31$$

such that

$$J(\omega) = \overline{B_{local}^2} j(\omega) \qquad 4.32$$

Therefore at the Larmor frequency, the reduced spectral density is described by

$$j(\omega_0) = \frac{2\tau_c}{1 + \omega_0^2\tau_c^2} \qquad 4.33$$

In the fast motion limit, τ_c is very small (e.g. in the case for small molecules), $\omega_0\tau_c \ll 1$, therefore $j(\omega_0) = 2\tau_c$, i.e. the reduced spectral density function is independent of ω_0.

In the slow motion limit, (e.g. in the case of small proteins), $\omega_0\tau_c \gg 1$, therefore $j(\omega_0) = \frac{2}{\omega_0^2\tau_c}$.

The rate constant for longitudinal relaxation, R_z, is given by

$$R_z = \gamma^2 \overline{B_{local}^2} j(\omega_0) \quad 4.34$$

i.e. R_z is dependent on the reduced spectral density function.

The rate constant for transverse relaxation, R_{xy}, is given by

$$R_{xy} = \frac{1}{2}\boxed{\gamma^2 \overline{B_{local}^2} j(0)} + \frac{1}{2}\boxed{\gamma^2 \overline{B_{local}^2} j(\omega_0)} \quad 4.35$$

secular non-secular

i.e. it depends on both $j(0)$ (the maximum value of the reduced spectral density function when $\omega = 0$) and $j(\omega_0)$. Importantly, the non-secular term is equal to $\frac{1}{2}R_z$, i.e. the longitudinal rate constant, R_z, contributes to the transverse rate relaxation, R_{xy}.

At fast motion, $j(\omega_0) = 2\tau_c$ and $j(0) = 2\tau_c$ so the relaxation rate constants can be re-written as

$$R_z = 2\gamma^2 \overline{B_{local}^2} \tau_c \quad 4.36$$

$$R_{xy} = 2\gamma^2 \overline{B_{local}^2} \tau_c \quad 4.37$$

This means that for small molecules, $T_1 = T_2$, i.e. magnetisation disappears from the xy-plane and re-appears on the z-axis at the same rate.

At slow motion, $j(0) = \frac{2}{\omega_0^2 \tau_c}$, so the relaxation rate constants can be re-written as

$$R_z = \frac{2\gamma^2 \overline{B_{local}^2}}{\omega_0^2 \tau_c} \quad 4.38$$

$$R_{xy} = \gamma^2 \overline{B_{local}^2} \tau_c \quad 4.39$$

This means that for large molecules $T_1 \gg T_2$, i.e. magnetisation disappears from the xy-plane long before it re-appears on the z-axis.

Finally, the $\frac{T1}{T2}$ ratio, can be used to report on the apparent correlation time for backbone amides. Low values are indicative of local, highly dynamic motions on the ps-ns time-scale, while high values are indicative of possible conformational changes on the ms time-scale.

In practice, for ^{15}N T_1 and T_2 experiments, ^{15}N-HSQC experiments are recorded with a series of increasing delays during which relaxation occurs. Peak intensities can then be fitted to an exponential decay

$$I(t) = I_0 e^{-Rt} \quad 4.40$$

where $I(t)$ is peak intensity at delay t, I_0 is peak intensity when the delay is zero and R is the rate of relaxation.

4.5.1.2 The steady-state NOE experiment

In steady-state NOE relaxation analysis (Keeler, 2005), a weak field is used to irradiate and therefore saturate a particular spin, i.e. z-magnetisation becomes equal to zero, e.g. when magnetisation of spin two is saturated, $\hat{I}_{2z} = 0$.

Irradiating spin two long enough leads to a steady state (ss) where magnetisation of spin one is constant and does not change with time, which can be written as

$$\left(\frac{d\hat{I}_{1z}}{dt}\right)_{ss} = 0 \quad 4.41$$

The Solomon equation for spin $\hat{I}_{1z}$ magnetisation is expressed as

$$\frac{d\hat{I}_{1z}}{dt} = -R_z^{(1)}\left(\hat{I}_{1z} - \hat{I}_{1z}^0\right) - \sigma_{12}\left(\hat{I}_{2z} - \hat{I}_{2z}^0\right) \quad 4.42$$

where $R_z^{(1)}$ is the rate of self-relaxation of spin one, $\hat{I}_{1z}$ is the magnetisation of spin one, $\hat{I}_{1z}^0$ is the spin one magnetisation at equilibrium, $\hat{I}_{2z}$ is the magnetisation of spin two, $\hat{I}_{2z}^0$ is the spin two magnetisation at equilibrium and σ_{12} is the rate of cross-relaxation.

At steady state conditions, $\hat{I}_{2z} = 0$ such that

$$\left(\frac{d\hat{I}_{1z}}{dt}\right)_{ss} = 0 = -R_z^{(1)}\left(\hat{I}_{1z,ss} - \hat{I}_{1z}^0\right) - \sigma_{12}\left(-\hat{I}_{2z}^0\right) \quad 4.43$$

$$0 = -R_z^{(I)}\left(\hat{I}_{1z,ss} - \hat{I}_{1z}^0\right) - \sigma_{IS}\left(-\hat{I}_{2z}^0\right) \quad 4.44$$

Rearranging for the steady-state value of z-magnetisation on spin $\hat{I}_{1z,ss}$, gives

$$\hat{I}_{1z,ss} = \frac{\sigma_{12}}{R_z^{(1)}}\hat{I}_{2z}^0 + \hat{I}_{1z}^0 \qquad 4.45$$

which shows that cross-relaxation and the longitudinal relaxation rate of spin one both control the steady state NOE.

The NOE enhancement by this experiment, η_{ss}, is determined by

$$\eta_{ss} = \frac{\left(\frac{\sigma_{12}}{R_z^{(1)}}\hat{I}_{2z}^0 + \hat{I}_{1z}^0\right) - \hat{I}_{1z}^0}{\hat{I}_{1z}^0} = \frac{\sigma_{12}}{R_z^{(1)}} \qquad 4.46$$

on the assumption that $\hat{I}_{1z}^0 = \hat{I}_{2z}^0$.

4.5.1.3 The heteronuclear steady-state NOE experiment

In the heteronuclear steady-state NOE experiment, spin one is the proton ($\hat{I}$) and spin two is the heteronucleus ($\hat{S}$), e.g. ^{15}N. The ^{1}H is saturated and changes to the ^{15}N signal are observed. The experiment consists of a single re-focused INEPT where the magnetisation at the start is on the ^{15}N and is subsequently transferred onto the ^{1}H.

The steady-state value of z-magnetisation on the heteronuclear spin, $\hat{S}_{z,ss}$, is given by

$$\hat{S}_{z,ss} = \frac{\sigma_{IS} I_z^0}{R_z^{(S)}} + \hat{S}_z^0 \qquad 4.47$$

where $R_z^{(S)}$ is the rate of self-relaxation, $\hat{I}_z^0$ is the spin $\hat{I}$ magnetisation at equilibrium, $\hat{S}_z^0$ is the spin $\hat{S}$ magnetisation at equilibrium and σ_{IS} is the rate of cross-relaxation.

The ratio of the equilibrium spin populations of the two spins is the same as the ratio of the gyromagnetic ratios of those spins

$$\frac{\hat{S}_z^0}{\hat{I}_z^0} = \frac{\gamma_S}{\gamma_I} \qquad 4.48$$

where $\gamma_I = 267 \times 10^6\ rad\ s^{-1}\ T^{-1}$ (for ^{1}H) and $\gamma_S = -27 \times 10^6\ rad\ s^{-1}\ T^{-1}$ (for ^{15}N).

Rearranging for $\hat{I}_z^0$ gives

$$\hat{I}_z^0 = \frac{\hat{S}_z^0 \gamma_I}{\gamma_S} \quad 4.49$$

which when substituted in the expression of $\hat{S}_{z,ss}$ gives

$$\hat{S}_{z,ss} = \frac{\gamma_I \sigma_{IS} I_z^0}{\gamma_S R_z^{(S)}} + \hat{S}_z^0 \quad 4.50$$

Therefore the NOE enhancement by the heteronuclear steady-state NOE experiment, $\eta_{S,ss}$, is given by

$$\eta_{S,ss} = \frac{\gamma_I \sigma_{IS}}{\gamma_S R_z^{(S)}} \quad 4.51$$

In the limit of fast motion, when $\omega_0 \tau_c \ll 1$, the reduced spectral density function is:

$$j(\omega_0) = \frac{2\tau_c}{1 + \omega_0^2 \tau_c^2} = 2\tau_c \quad 4.52$$

Therefore from the Solomon equations, we can get simplified expressions for $R_z^{(S)}$ and σ_{IS},

$$R_z^{(S)} = b^2 \tau_c \quad 4.53$$

$$\sigma_{IS} = \frac{b^2 \tau_c}{2} \quad 4.54$$

where

$$b = \frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi r^3} \quad 4.55$$

(r represents the distance between the two spins, μ_0 is a constant termed the permeability of vacuum with value $4\pi \times 10^{-7} H\ m^{-1}$ and ħ is Planck's constant).

Substituting the $R_z^{(S)}$ and σ_{IS} expressions into the expression for $\eta_{S,ss}$ simplifies it to

$$\eta_{S,ss} = \frac{\gamma_I}{2\gamma_S} \quad 4.56$$

for fast motion (in small organic molecules).

As γ_I (^{1}H) is positive and γ_S(^{15}N) is negative, $\eta_{S,ss}$ will be negative for spins in fast motion. This means that saturating the protons of highly dynamic backbone amides results in large,

negative ^{15}N signals, while for less dynamic backbone amides the $\eta_{S,ss}$ will be approximately zero.

In practice, in heteronuclear steady-state NOE relaxation analysis, separate spectra are recorded with and without proton saturation (saturated and reference spectra, respectively). Heteronuclear NOE values are established by calculating the ratio of peak intensities

$$NOE = \frac{I_{saturated}}{I_{reference}} \qquad 4.57$$

where I represents peak intensity.

4.5.2 ^{15}N-relaxation experiments on PRK1 HR1c

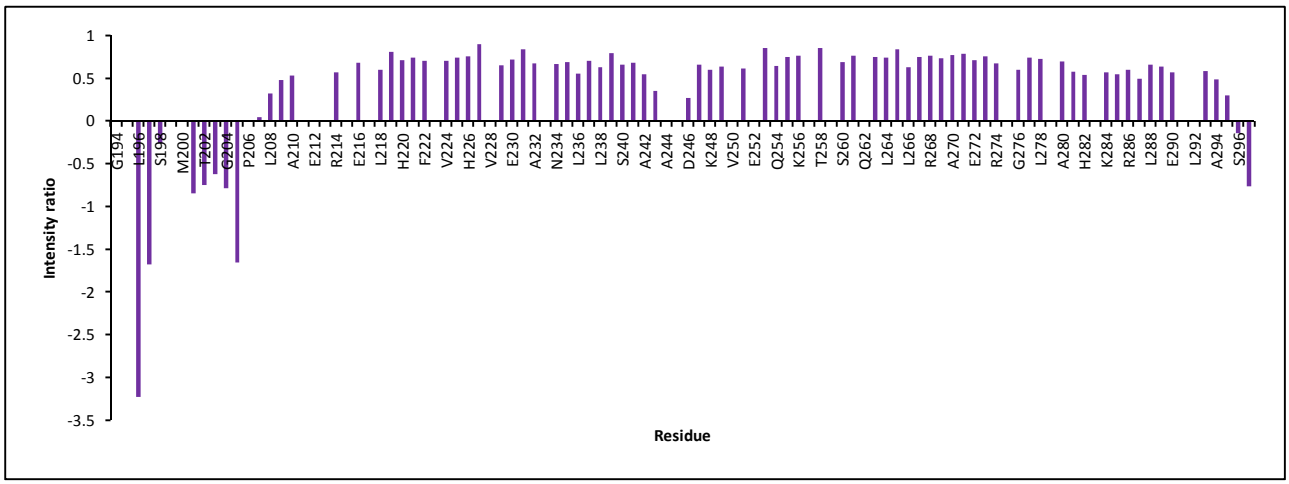
A heteronuclear steady-state NOE experiment was recorded on PRK1 HR1c and intensity ratios were calculated (Figure 4.29A). Negative values were obtained for N-terminal residues as well as the two C-terminal residues, which are unstructured and are therefore highly dynamic. A smaller intensity ratio was also seen with Ala242 and to a much greater extent Lys243 and Asp246, which are in the inter-helical loop. Similarly smaller intensity ratios were seen for residues in the second inter-helical loop (Ala280-His282). In contrast, the intensity ratios for the residues in all three helices were positive, suggesting that helix 3 is truly fixed on the other two helices and is not swinging around in solution.

T_1 and T_2 constant times were also obtained (Figures 4.29B-C). For helices 1 and 2, T_1 values of most residues were greater than 0.8 s. In contrast, for helix 3 residues, T_1 values ranged from 0.703-0.836 s, while T_1 values for residues in unstructured regions were typically lower (~0.6 s). T_2 values of most residues in all three helices and both inter-helical loops were ~0.05 s, while for the unstructured N-terminus and the last two C-terminal residues, they were greater, with the greatest T_2 value being ~15 times larger than the smallest T_2 value.

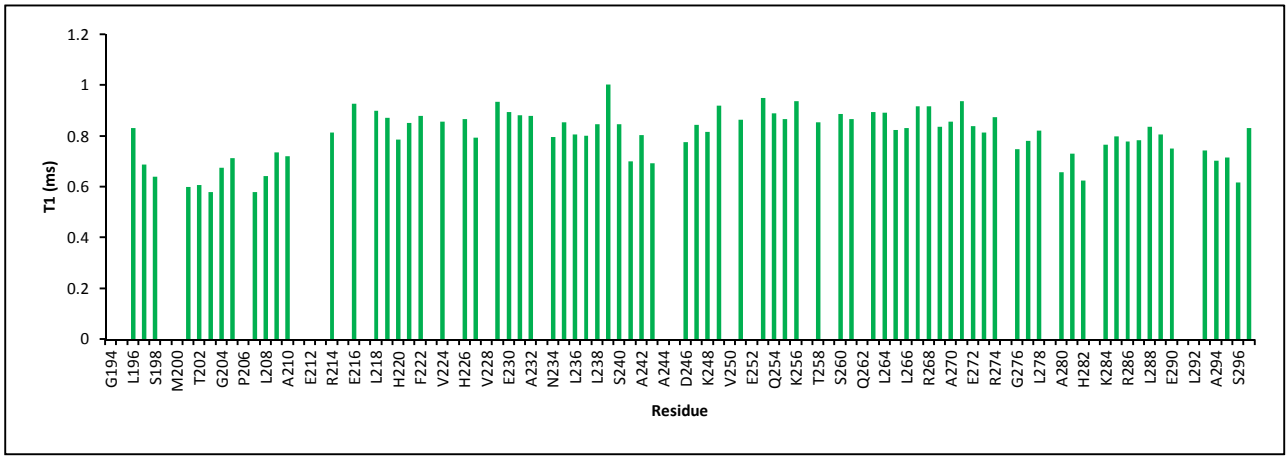
An interesting observation was made when then $\frac{T_1}{T_2}$ ratio was plotted. The average ratio across helix 1 was 17.5 while the average ratio across helix 2 was 18.4. In contrast, the

average ratio across helix 3 was 13.4, suggesting that this helix has a different orientation with respect to the principal diffusion axis of the molecule (the coiled coil), i.e. it is not completely parallel to the principal axis, which is consistent with the structure ensemble.

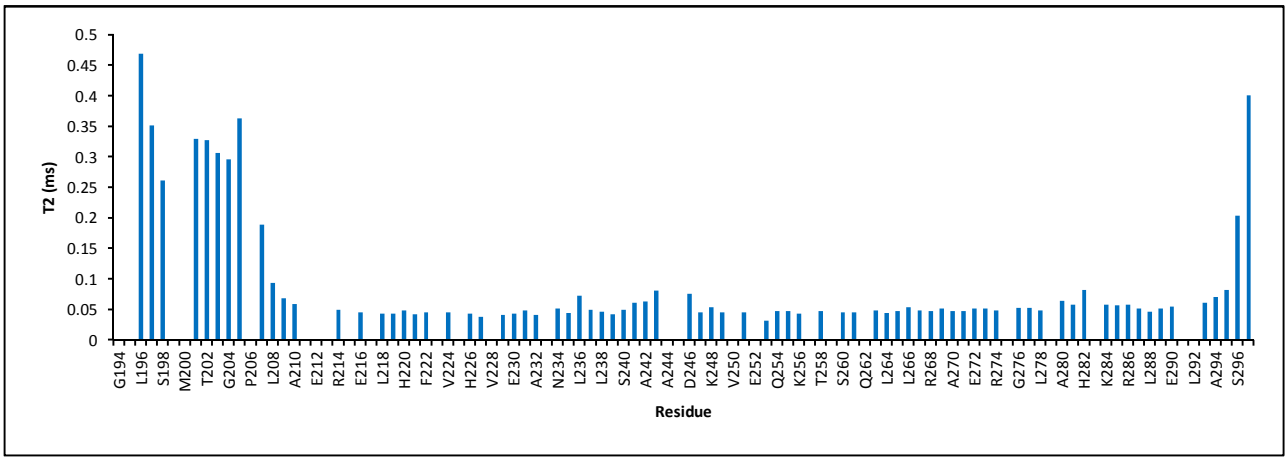
A


Intensity ratio
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
G194 L196 S198 M200 T202 G204 P206 L208 A210 E212 R214 E216 L218 H220 F222 V224 H226 V228 E230 A232 N234 L236 L238 S240 A242 A244 D246 K248 V250 E252 Q254 K256 T258 S260 Q262 L264 L266 R268 A270 E272 R274 G276 L278 A280 H282 K284 R286 L288 E290 L292 A294 S296
Residue

B


T1 (ms)
1.2
1
0.8
0.6
0.4
0.2
0
G194 L196 S198 M200 T202 G204 P206 L208 A210 E212 R214 E216 L218 H220 F222 V224 H226 V228 E230 A232 N234 L236 L238 S240 A242 A244 D246 K248 V250 E252 Q254 K256 T258 S260 Q262 L264 L266 R268 A270 E272 R274 G276 L278 A280 H282 K284 R286 L288 E290 L292 A294 S296
Residue

C


T2 (ms)
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
G194 L196 S198 M200 T202 G204 P206 L208 A210 E212 R214 E216 L218 H220 F222 V224 H226 V228 E230 A232 N234 L236 L238 S240 A242 A244 D246 K248 V250 E252 Q254 K256 T258 S260 Q262 L264 L266 R268 A270 E272 R274 G276 L278 A280 H282 K284 R286 L288 E290 L292 A294 S296
Residue

D

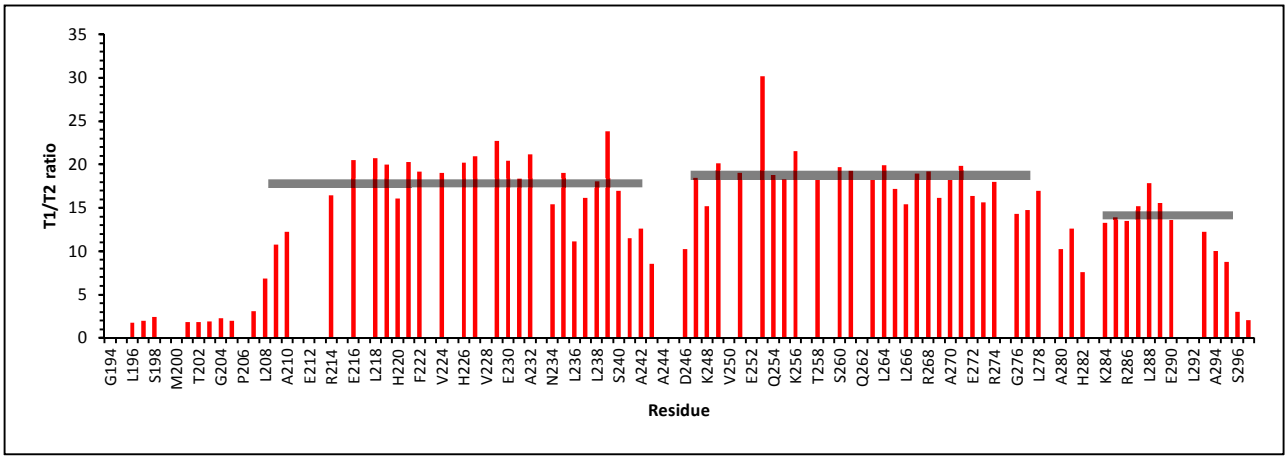


Figure 4.29 ^{15}N relaxation measurements for PRK1 HR1c. Experiments were recorded with a 1.6 mM PRK1 HR1c sample in MTPBS pH 7.3 and 10% D_2O recorded on a Bruker DRX500 at 25 ^{o}C. Values were obtained from peak intensities from the relevant ^{15}N-HSQC spectra. (**A**) ^{1}H-^{15}N heteronuclear NOE intensity ratios (saturated to reference experiment) for each residue. (**B**) T_1 constant values for each residue. (**C**) T_2 constant values for each residue. (**D**) T_1/T_2 ratio for each residue. The grey bars indicate the mean T_1/T_2 ratio across each helix.

4.6 Investigating the structure of PRK3 HR1c

4.6.1 Sequence alignment of the two HR1c domains

The HR1c domains from PRK1 and PRK3 have high sequence identity and sequence similarity (Figure 4.30), yet it has been shown in Chapter 3 that they have considerable differences in terms of their oligomerisation and stability. PRK1 HR1c is monomeric in solution and has a melting temperature which is 5 ^{o}C higher than that of PRK3 HR1c, which oligomerises in solution. The solution structure of PRK1 HR1c confirmed secondary structure predictions and showed that both the two main helices but also the C-terminal helix actually exist. The differences between these two domains however suggest that PRK3 HR1c may structurally differ. Leu292^{PRK1} (helix 3) is conserved as Leu257^{PRK3}, suggesting that the C-terminal helix may be packed against the other two helices of PRK3 HR1c. While Phe222^{PRK1} is strongly conserved as Leu187^{PRK3}, this may prevent formation of the triple coiled coil, as the phenyl ring is crucial for several interactions with helix 2 and helix 3 side-chains.

Interestingly, a leucine, a glutamate and an alanine in the C-terminal helix of PRK1 are replaced by three arginine residues in PRK3 which are on the same side of the helix, spanning two consecutive helical turns. These residues are solvent exposed in the PRK1 HR1c structure but in PRK3 HR1c they may be involved in electrostatic interactions, which could possibly drive PRK3 HR1c oligomerisation. The C-terminal helix in PRK3 HR1c is the least conserved helix which might explain the significant difference in the oligomeric state of these HR1 domains.

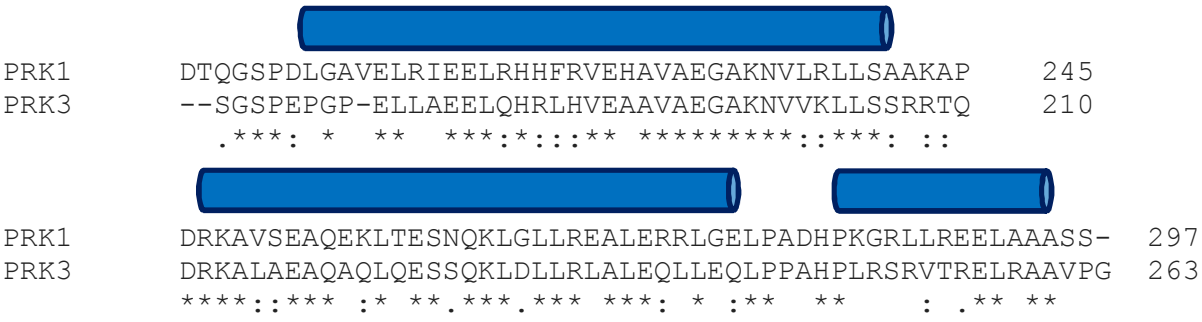


Figure 4.30 Sequence alignment of PRK1 HR1c and PRK3 HR1c. Sequences were aligned with Clustal Omega. Asterisks highlight identical residues at all positions, colons highlight conservation of a strong group and periods highlight conservation of a weak group. The cylinders represent α-helices based on the structure of PRK1 HR1c.

4.6.2 Investigating the structure of PRK3 HR1c by NMR

To explore the possibility of studying the structure of PRK3 HR1c by NMR, the quality of NMR spectra of this protein was assessed by recording 2D-NOESY and 2D-TOCSY spectra on unlabelled protein. In contrast to the strong, sharp peaks seen in the 2D-NOESY and 2D-TOCSY of PRK1 HR1c, weak peaks were observed in the case of PRK3 HR1c (Figure 4.31). This is a result of protein oligomerisation where there is dynamic exchange between the different oligomeric states, leading to broadening of the NMR signals. This is consistent with the AUC findings from Chapter 3. The structure of PRK3 HR1c was therefore not investigated any further by NMR.

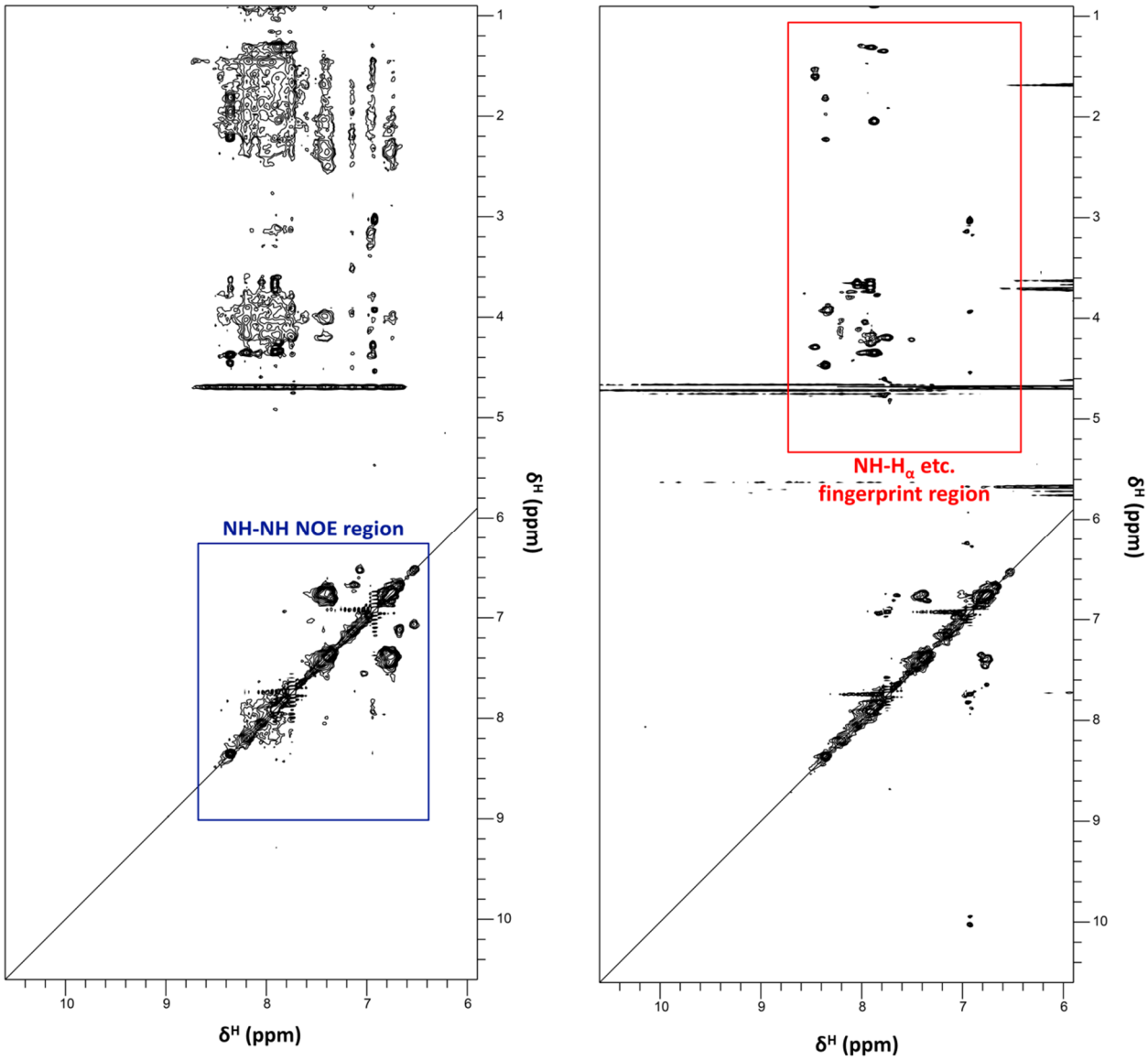


Figure 4.31 2D NOESY and 2D TOCSY spectra of PRK3 HR1c. NMR spectra of 0.7 mM PRK3 HR1c in 20 mM sodium phosphate pH 7.3, 150 mM NaCl and 10% D_2O were recorded on a Bruker DRX500 at 25 ^{o}C. The amide region which contains NH-NH NOEs is highlighted in the 2D NOESY while the NH-H_α etc. fingerprint region is highlighted in the 2D TOCSY.

4.7 Conclusions

This chapter details a solution NMR approach in determining the structure of PRK1 HR1c. The structure shows that HR1c consists of three helices. The two longest helices form an antiparallel coiled coil which is analogous to that of HR1a and HR1b. In addition, the third helix packs against the other two helices to form a triple helix. This C-terminal helix is absent in both HR1a and HR1b so it is a unique feature of the HR1c domain. The functional role of this helix remains unknown but it may have a role in the interactions of this HR1 domain with Rho GTPases.

Similarly to HR1a and HR1b, HR1c is likely to also be a type D anti-parallel coiled coil effector of GTPases. Type D effectors are the most popular type of helical GTPase effector and include effectors of Rho, Rab and Arf GTPases (reviewed in Mott & Owen, 2015). Effectors of this type interact with the GTPase via a longer N-terminal helix and a shorter C-terminal helix. While the majority of helical GTPase effectors use a single pair of helices to bind to the GTPase, there are some exceptions. For example, mDia1 (another type D effector) forms two helices which interact with the switch regions of RhoC in a GTP-dependent manner (Rose *et al.*, 2005). A third helix packs against these helices and stabilises their orientation, promoting GTPase binding. This stabilising helix also contributes two residues to the interaction with switch 2 of RhoC. The C-terminal helix of HR1c, which also packs against the two main helices, may therefore have a similar role in GTPase binding.

The lack of published interacting partners of HR1c may suggest that this domain evolved away from the other HR1 domains and does not bind to GTPases. However HR1c may instead be involved in inter-domain interactions (either intra-molecular or inter-molecular) that are crucial for the formation of the HR1abc trimer (see Chapter 3). The C-terminal helix may have a role in mediating these interactions, which could explain why it is unique to the HR1c domain. The solution structure of HR1c is a powerful tool in the pursuit of structural characterisation of inter-domain interactions, which were investigated in the next chapter.

Chapter 5 Investigating the intra- and inter-molecular interactions between HR1 domains

5.1 Introduction

A key question regarding PRK1 is whether the three HR1 domains interact with each other or exist independently in solution. As shown in Chapter 3, HR1b and HR1c are monomeric in solution while HR1a and the HR1ab di-domain dimerise, suggesting that PRK1 dimerisation is mediated by inter-molecular HR1a-HR1a interactions. The characterisation of HR1abc by AUC however indicated that the tri-domain exists as monomer, dimer and trimer in solution. The trimer is indicative of further interactions in the context of HR1abc which only occur when HR1c is present in the context of HR1a and HR1b. These interactions could occur intra-molecularly, i.e. between HR1 domains of a single HR1abc protomer, or inter-molecularly, i.e. between two or more protomers, driving trimer formation. Such interactions may be significant for the regulation of PRK1 activity, e.g. formation of inactive oligomers as seen in PRK2 (Bauer *et al.*, 2012) or differential binding of oligomers to GTPases. The aim of this Chapter will therefore be to investigate these interactions and to structurally characterise HR1abc.

5.2 Investigating the structure of the HR1abc tri-domain

The gold standard to elucidate the intra-molecular and inter-molecular interactions is to obtain a structure of HR1abc. The following sections outline attempts to do this by NMR spectroscopy and X-ray crystallography.

5.2.1 Investigating HR1abc by NMR

To determine whether the quality of HR1abc spectra would enable the structural characterisation of HR1abc by solution NMR, a 2D NOESY spectrum was recorded on an unlabelled sample at 20 ^{o}C. The resulting spectrum contained few NH-NH and NH-H_{α} NOEs due to significant NMR signal linewidth broadening. Despite the fact that the spectrum was recorded at a concentration of 450 μM, where the trimer is the predominant species in solution (73.7% trimer), the spectrum was of poor quality. This is the result of the chemical exchange in the intermediate regime between the oligomeric states of HR1abc as observed previously in AUC experiments.

To investigate whether the chemical exchange could sufficiently be decreased, a 2D NOESY was recorded at 5 ^{o}C. Figure 5.1 shows the overlaid 2D NOESY temperature titration spectra. As expected, at the lower temperature there was significantly stronger signal strength, as well as improved resolution. A number of new NOEs also appeared, suggesting a reduced rate of chemical exchange between the various oligomeric states of HR1abc. However, in comparison to other 2D NOESY spectra, such as that of HR1c (Chapter 4), there is still significant peak broadening and many NOEs are not visible. As a result, further investigation of the HR1abc structure by NMR was not possible.

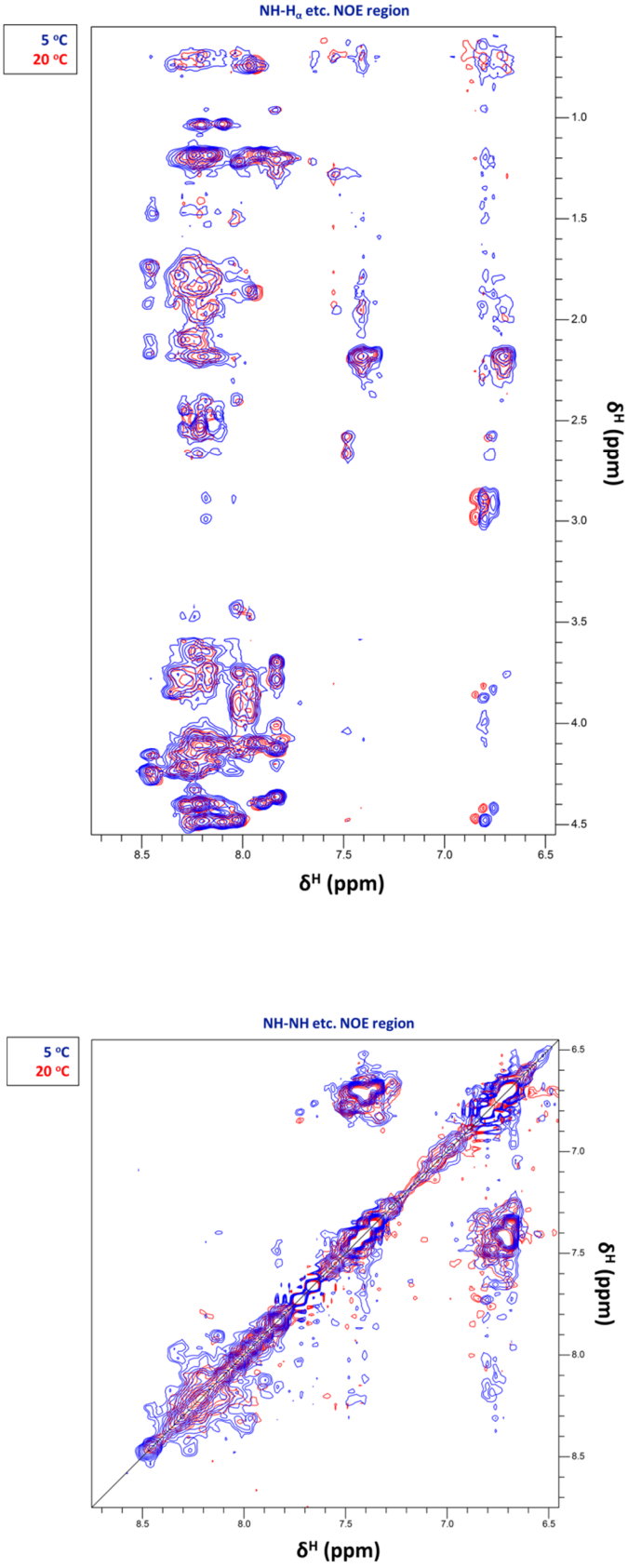


Figure 5.1 2D NOESY titration spectra of PRK1 HR1abc. Overlaid NMR spectra of 0.45 mM PRK1 HR1abc in 20 mM sodium phosphate pH 7.3, 150 mM NaCl, 5 mM DTT and 10% D_2O were recorded on a Bruker DRX500 at 5 ^{o}C (blue) and 20 ^{o}C (red).

5.2.2 Crystallisation trials

5.2.2.1 HR1abc crystallisation trials

Initial crystallisation trials were set up in crystallisation plates screening a wide range of compounds. Drops were set up by mixing PRK1 HR1abc at 10 mg/ml at 1:1 and 1:0.5 ratios with the buffer in each well. Some phase separation was seen as early as one week but no further changes were observed in the next two months. Phase separation was primarily observed in wells with polyethylene glycol 3350 or polyethylene glycol 4000 as one of the components (Figure 5.2A).

A second round of crystallisation trials was then carried out in order to screen around these conditions and therefore maximise the probability that a crystal would form. While phase separation was observed, as expected, no crystals formed even after one year. This suggests that the polydispersity of HR1abc opposes crystal packing.

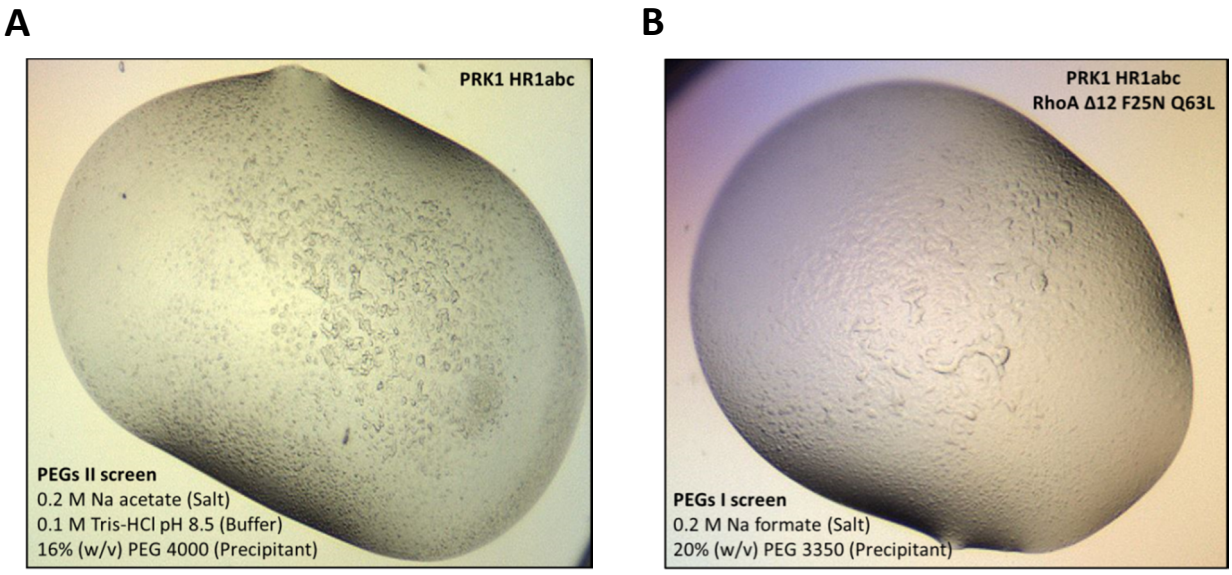


Figure 5.2 PRK1 HR1abc crystallisation trials. (**A**) Example of drop with phase separation as seen in PRK1 HR1abc crystallisation trials set up with 10 mg/ml PRK1 HR1abc. (**B**) Example of drop with phase separation as seen in PRK1 HR1abc/RhoA Δ12 F25N Q63L co-crystallisation trials set up with 250 µM PRK1 HR1abc and 250 µM RhoA Δ12 F25N Q63L. The screen used and the buffer components are indicated.

5.2.2.2 HR1abc + RhoA Δ12 F25N Q63L crystallisation trials

Since the crystal structure of HR1a was solved as a complex with RhoA (Maesaki *et al.*, 1999), it is possible that RhoA stabilised HR1a and promoted crystal formation. Therefore a similar approach was also employed with HR1abc. RhoA was purified from *E. coli* from a construct lacking the unstructured C-terminal tail which would hinder ordered crystal packing (Δ12). The F25N mutation makes recombinant RhoA soluble with no reported functional implications and it has been widely used in biochemical and biophysical studies. The Q63L mutation was used to 'lock' the GTPase in a constitutively active conformation. RhoA Δ12 F25N Q63L was purified and loaded with the non-hydrolysable analogue GMPPNP in the same way as RhoA Δ4 (see Chapter 6 for purification details). Crystallisation trials were set up using a 1:1 mixture of RhoA:HR1abc at a concentration of 250 mM for each protein. Some phase separation was observed during week one, however as with the previous crystallisation trials no crystals formed (Figure 5.2B). It therefore appeared that obtaining the structure of HR1abc was not amenable by X-ray crystallography.

It is worth nothing that further screening could have been carried out at different protein concentrations, lower temperatures or using other buffers. However given the polydispersity of HR1abc these attempts would also probably have been unsuccessful.

5.3 Investigating the interactions between the individual HR1 domains

5.3.1 Rationale

Section 5.2 highlights the experimental obstacles encountered when attempting to investigate the structure of the HR1abc tri-domain. In brevity, the reversible intra- and inter-molecular interactions of HR1abc prevent its structural characterisation by NMR and X-ray crystallography. Without a structure it is difficult to know whether the HR1 domains interact or are independent of each other but also, if they interact, which contacts they form and how. Structural studies of equivalent systems have been published but they are driven by mutations that stabilise the protein in a single oligomeric state. Therefore a monodisperse sample can be obtained and structurally characterised, e.g. as with amyloid-β oligomers (Sandberg *et al.*, 2010). This was not possible with HR1abc as the HR1a dimerisation site and inter-HR1 contact sites are not known. In addition, even if the structure of such a monomeric mutant protein was available, it may not be physiologically relevant.

To bypass this problem, it was decided to dissect the sum of intra- and inter-molecular interactions by studying interactions between individual HR1 domains or between the HR1ab di-domain and HR1c. The aim was to obtain different puzzle pieces which would be put together to give us a structural understanding of HR1abc.

5.3.2 Isothermal Titration Calorimetry experiments

Isothermal Titration Calorimetry (ITC) is a biophysical method used to study protein-protein interactions. It involves heat changes occurring when two proteins interact. These heat changes are produced when a protein in a syringe is gradually injected into another protein in a cell. This allows the derivation of parameters such as K_d, entropy change and stoichiometry of the interaction. ITC depends on polar or charged residues involved in the interaction, so it is not useful in identifying or characterising protein-protein interactions which are driven solely by hydrophobic contacts.

If the HR1 domains interact via charged or polar residues, then a heat change is expected upon such an interaction. ITC was used to not only identify such HR1 inter-domain interactions, but also to provide a measure of the affinity between the two domains. While these affinities would be of isolated domains *in vitro*, they would still be helpful, e.g. to compare to published affinities of HR1 domains for GTPases. This could help understand if GTPases could potentially disrupt these interactions.

In this case, the HR1ab di-domain was placed in the cell and HR1c was titrated from the syringe. Since HR1ab can dimerise, it would be disadvantageous to place it in the syringe, as the change from high to low concentration upon injection could lead to release of a heat of dilution which would complicate the interpretation and potentially obscure any heat changes arising from HR1ab interacting with HR1c. The experiment was repeated at two concentrations; 500 µM HR1c into 50 µM HR1ab and 5 mM HR1c into 500 µM HR1ab. No heat changes were observed in either case, indicating that either HR1ab does not interact with HR1c or that any inter-domain interactions must be solely hydrophobic.

5.3.3 NMR titrations with individual domains

The HR1abc trimer, the formation of which is thought to be driven by intra-molecular and inter-molecular interactions, becomes the primary species in solution only at high µM concentrations (above 400 µM; see Chapter 3). This suggests that any possible interactions between HR1 domains are weak, making them difficult to be detected by techniques such as pull-downs using recombinant proteins. This is because the HR1 proteins would need to be at concentrations at least ten-fold above the K_d (to ensure binding) and this is not possible. If these interactions are solely hydrophobic, then methods such as ITC also fail, as discussed in section 5.3.2. Therefore NMR was used to investigate these potentially weak, hydrophobic interactions.

5.3.3.1 Theory

NMR is a powerful tool for studying protein-protein interactions. The sensitivity of NMR enables the fingerprinting of potentially all of the individual protons in the protein. The ability of NMR to detect tiny changes in the chemical environment of protons means that it can be employed to study all kinds of protein-protein interactions, including transient, weak interactions. Typically, NMR protein-protein interaction studies are carried out as titrations, where an unlabelled protein is gradually titrated into a labelled protein and experiments, e.g. ^{15}N HSQC, are recorded at each titration point. Interactions are detected in the form of chemical shift perturbations (CSPs) which reflect a change in the chemical environment of protons.

Chemical exchange refers to the dynamic interconversion of a proton between two or more distinct states or environments (reviewed in Kleckner & Foster, 2010; Williamson, 2013). NMR spectra can be affected by chemical exchange and are typically categorised in three exchange regimes (Figure 5.3) based on the rate of the chemical exchange, k_{ex}, which is a measure of the stochastic exchange events per second (and is temperature-dependent). These regimes are slow exchange, intermediate exchange and fast exchange and are defined by the value of k_{ex} relative to the chemical shift difference, $\Delta\omega$.

For residues undergoing slow exchange ($k_{ex} \ll |\Delta\omega|$), two distinct peaks are observable. One peak is at the chemical shift of the original, unbound state and the second peak is at the chemical shift of the bound state. As the protein population in the bound state increases and the population of the free state decreases, the intensity of peaks of the bound state increases while the intensity of peaks of the unbound state decreases.

When $k_{ex} \gg |\Delta\omega|$, fast exchange occurs. A single peak is observed which is representative of both bound and unbound species populations. The peak is a population-weighted average and its exact chemical shift depends on the proportion of each of the bound and unbound species. Comparing such peaks from experiments with different ligand concentrations appears as if the peak is moving in a single direction. Fast exchange is often observed for protein-protein interactions with low affinities.

Intermediate exchange, as the name suggests, results when $k_{ex} \approx |\Delta\omega|$. In this regime, the peak is observed at an average chemical shift of the unbound and bound states.

Interconversion between bound and unbound states (ms timescale) during the detection step of NMR experiments leads to 'exchange broadening', i.e. peaks have broadened linewidths and may often 'disappear' from spectra (reviewed in Kleckner & Foster, 2010).

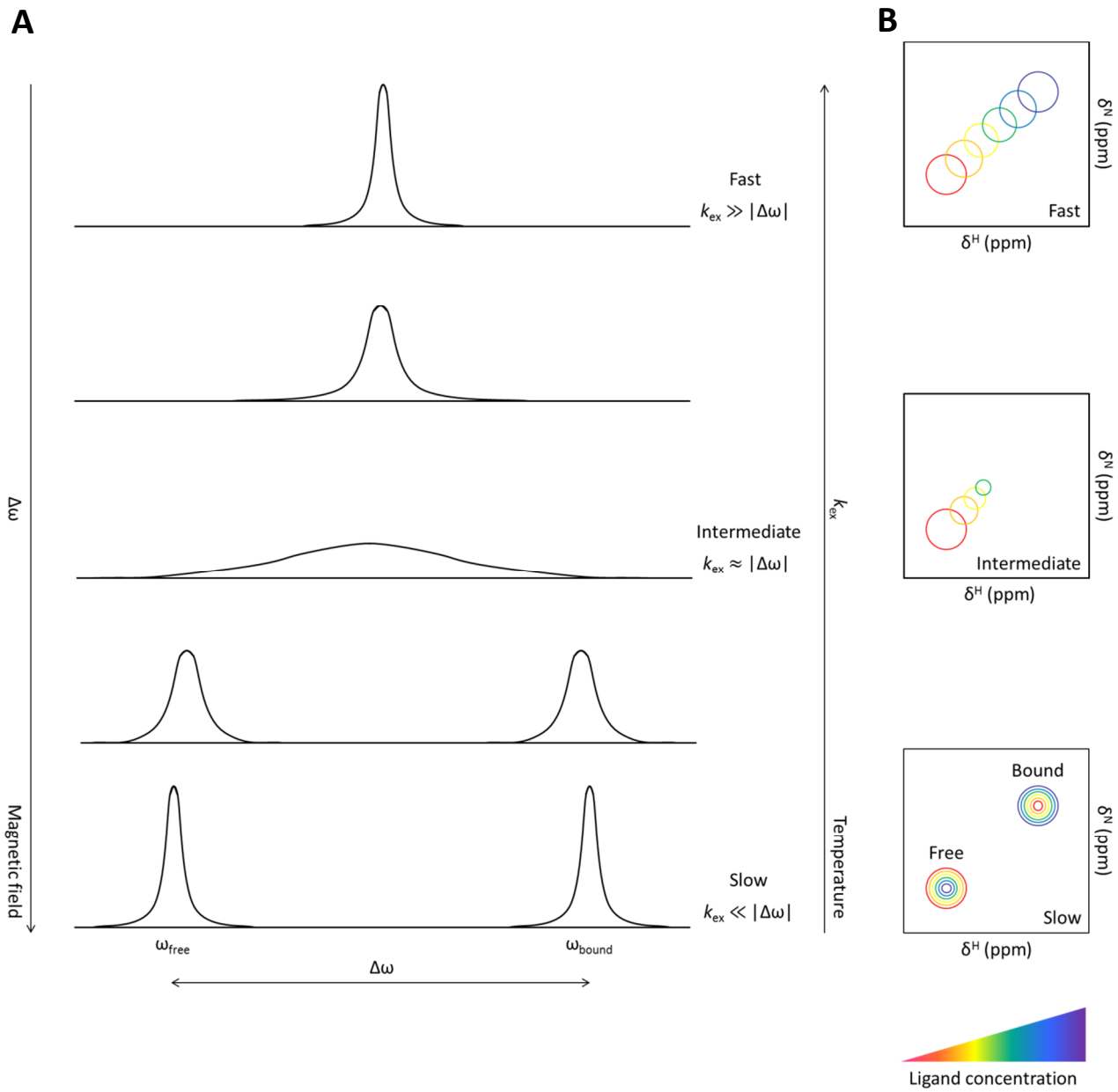


Figure 5.3 Chemical exchange regimes in NMR. (**A**) Examples of slow, intermediate and fast exchange of a single resonance in a 1D NMR spectrum. In the slow exchange regime, two peaks are observed, each at the chemical shift of the free (ω_{free}) and bound (ω_{bound}) states. In the fast exchange regime, a single peak is observed at a population-weighted average chemical shift of the free and bound states. Δω increases with a stronger magnetic field and k_{ex} increases with increasing temperature. In the intermediate exchange regime, the linewidth is increased as a result of exchange broadening. (**B**) Examples of slow, intermediate and fast exchange in a 2D spectrum. Peaks have the same shape throughout in the limit of fast exchange. In slow exchange, the intensity of the bound species peak increases with increasing ligand concentration, while the intensity of the free bound species decreases.

5.3.3.2 ^{15}N HR1c + HR1a

Unlabelled HR1a was titrated into ^{15}N HR1c and ^{15}N HSQC experiments were recorded. Small chemical shift changes were observed and shift changes continued to occur when the HR1c:HR1a ratio was increased from 1:2 to 1:6. Small chemical shift changes are indicative of a weak interaction between HR1c and HR1a. The chemical shift changes however are also small because the backbone amide protons are physically shielded in an α-helix. The shift changes were all in the fast exchange regime.

Two main binding sites were concluded from the data. These were indicated by the greatest shift changes across the entire protein (Figure 5.4). The first site, which is the one with the greatest shift changes, spans Val224 to Ala232. The exact sequence is VEHAVAEGA, which consists of a mixture of hydrophobic and charged residues. The greatest shift change occurs to Val228 at the centre of the sequence, which is solvent exposed (side-chain is 60% solvent exposed) according to NACCESS (Hubbard & Thornton, 1993). The largest shift changes occur with a periodicity of i to i+3, which is indicative of a single surface on the α-helix being used in the interaction. As with Val228, Glu225 and Gly231 are also solvent exposed. The remaining residues that shift, such as Ala229, probably do not contribute to the interaction with HR1a, but instead show chemical shift changes as a result of a propagated effect from the interacting residues.

A second site is also found in helix 1 of HR1c, spanning Val235 to Ala242, and Asp246 in the inter-helical loop. Helical i to i+3 periodicity is also observable here, specifically between Leu236, Leu239 and Ala242. Ala242 is the first residue of the flexible HR1c inter-helical loop (Ala242 to Asp246). The sequence of this second site is VLRLLSAAKAPD.

Smaller chemical shifts of approximately half the magnitude are observed with some helix 2 residues, specifically Gln254, Lys256, Thr258 and Ser260. However only Lys256 and Ser260 are on the same interface as the interacting residues above, therefore it is likely that if helix 2 of HR1c interacts with HR1a, it will be via these two residues.

It is important to keep in mind that HR1a also exists in a dynamic equilibrium between monomer and dimer as indicated by AUC experiments. This means that HR1a dimerisation and the HR1a interaction with HR1c are occurring simultaneously. It cannot be concluded from the data whether HR1c preferentially interacts with HR1a monomer or dimer. The

concentration of HR1a in these experiments ranges from 60 µM (1:0.25), where HR1a exists mainly as monomer, to 1.4 mM (1:6), where ~90% of HR1a is expected to be dimeric. The fact that the HR1a shift changes remain small even when a high concentration of HR1a is achieved can be explained in one of two ways. Firstly, that HR1c interacts with the HR1a dimer equally strongly to the monomer but the interaction with both is weak. Secondly, that HR1c interacts only with either the HR1a monomer or the dimer, i.e. the HR1c-HR1a interaction blocks or enhances the HR1a-HR1a interaction. Assuming that the HR1abc trimer is formed from the HR1abc dimer, it is unlikely that HR1a dimerisation is disrupted by the HR1a-HR1c interaction (which may drive trimer formation). Therefore HR1a dimerisation is unlikely to compete with the HR1a-HR1c interaction.

Plotting the normalised intensity changes for each residue reveals that there is an intensity drop from the beginning to the end-point of the titration. The intensity changes are not selective for any particular region of HR1c but are indicative of linewidth broadening due to formation of the HR1a-HR1c complex under intermediate exchange.

When mapping protein-protein interaction interfaces by NMR, titrations are often carried out reciprocally, using either ^{15}N-labelled protein. This was not done in this case as previous work in the lab has shown that the HSQC spectrum of ^{15}N HR1a looks broadened (due to the monomer-dimer exchange). Therefore it would not be possible to map the HR1c binding site on HR1a.

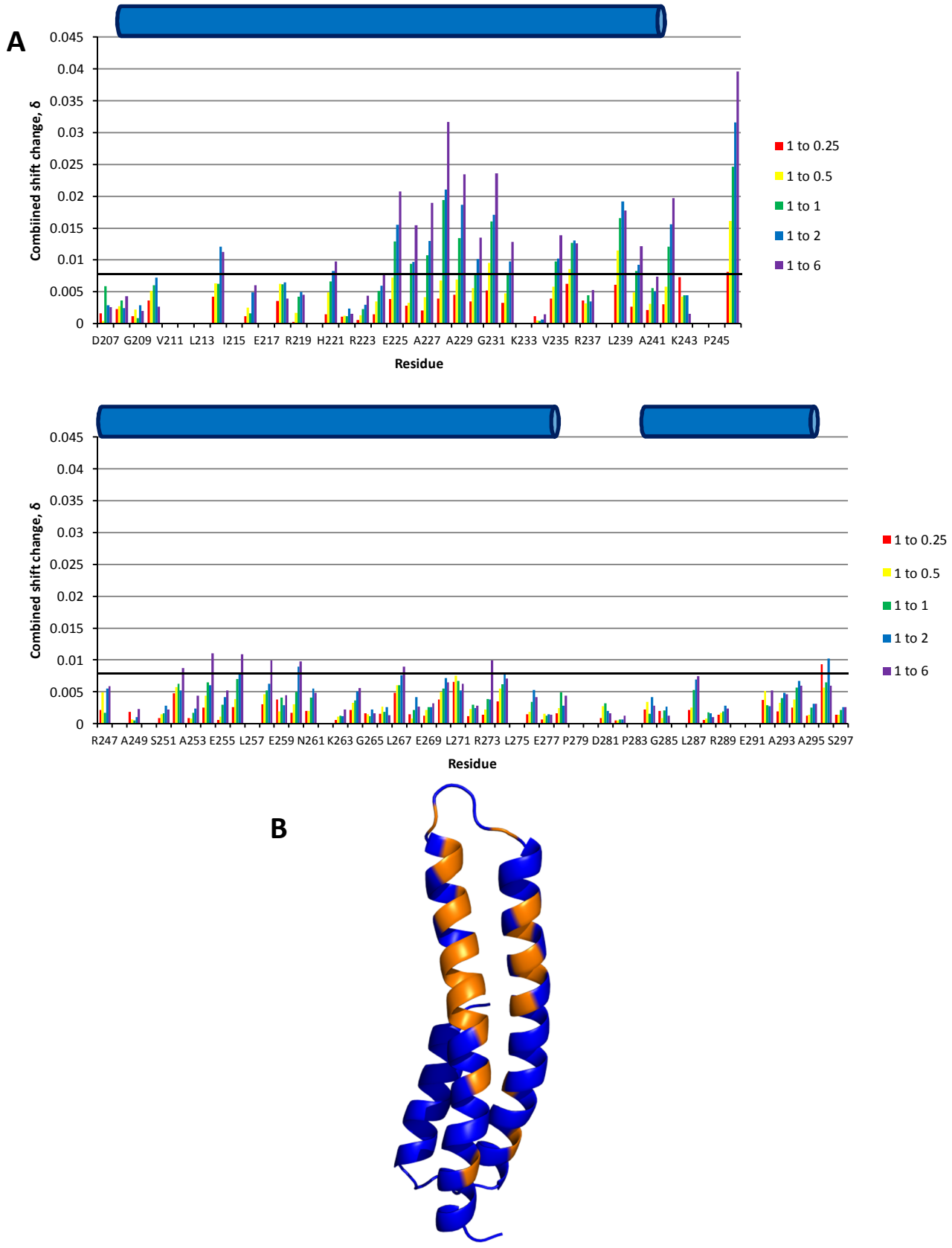


Figure 5.4 Quantitative analysis of ^{15}N HSQC spectra when HR1a was titrated into ^{15}N HR1c. (**A**) Bar chart showing the chemical shift changes for each residue of HR1c. The black line represents the mean combined shift change. Mean: 0.0078, standard deviation: 0.0074. (**B**) Residues which shifted more than the mean chemical shift change of all residues are coloured orange on the structure of HR1c.

5.3.3.3 ^{15}N HR1c + HR1b

To investigate whether HR1b also interacts with HR1c, HR1b was titrated into ^{15}N-labelled HR1c and HSQC experiments were recorded. Small chemical shift changes were observed and they were also indicative of fast exchange. The greatest shift changes observed were approximately half in magnitude compared to those in the HR1a-HR1c interaction (Figure 5.5). This suggests that HR1c interacts with HR1b more weakly than it does with HR1a.

One HR1b binding site appeared to be identical to the first HR1a binding site, i.e. VEHAVAEGA, with larger shift changes at i to i+3 intervals. Glu225, Val228 and Gly231 showed the biggest changes in this binding site, consistent with their solvent accessibility. This HR1c binding site is therefore shared by both HR1a and HR1b, suggesting that these two domains may compete for binding to the same site or that each one binds to this site in a different protomer in the context of an HR1abc oligomer. The second binding site was also shared with HR1a as indicated by the Val235, Leu239, A242 and D246 shift changes. A third potential binding site was seen in helix two and comprised Gln254, Thr258, Asn261, Leu264, Leu267 and Leu271, i.e. residues with helical periodicity, with the largest shift change being that of Leu267. While Gln254 and Thr258 also shifted in the HR1a titrations, the remaining residues (Asn261, Leu264, Leu267 and Leu271) had only shifted very slightly (section 5.3.3.2). Another difference with the HR1a titrations is that with HR1b the shift changes seen in helix 2 are of approximately the same magnitude as the shifts in helix 1, while in HR1a the few helix 2 shift changes were much smaller than the helix 1 shift changes. This suggests that both sites could equally be used by HR1c. Finally, non-selective intensity changes were also observed as with HR1a.

Interestingly, the HR1b contact sites are found on opposing sides of the HR1c coiled coil (Figure 5.5B). It is possible that one HR1b molecule binds to HR1c using contact sites one and two and another HR1b molecule binds to the same HR1c domain using contact site three.

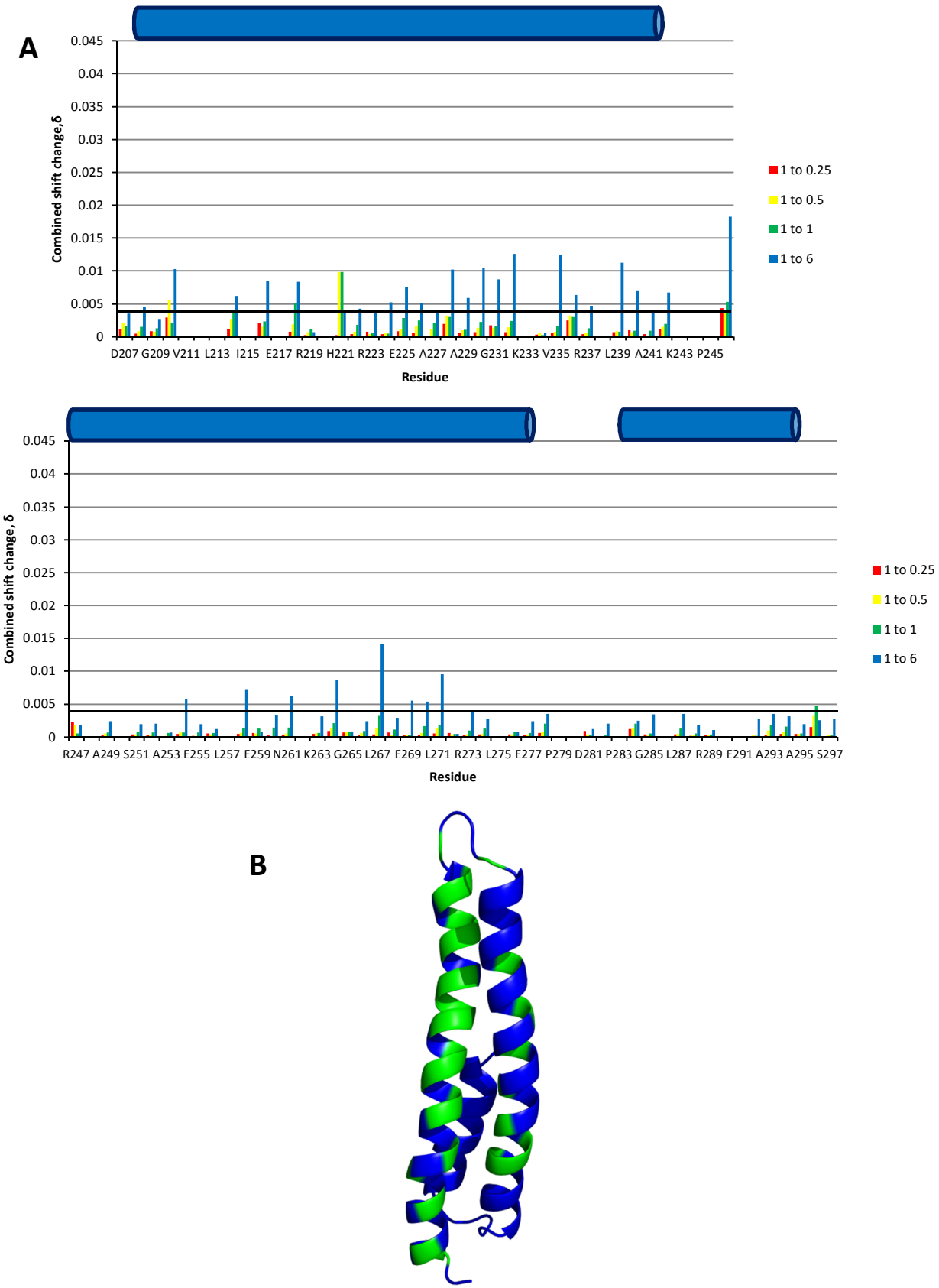


Figure 5.5 Quantitative analysis of ^{15}N HSQC spectra when HR1b was titrated into ^{15}N HR1c. (**A**) Bar chart showing the chemical shift changes for each residue of HR1c. The black line represents the mean combined shift change. Mean: 0.004, standard deviation: 0.0036. (**B**) Residues which shifted more than the mean chemical shift change of all residues are coloured green on the structure of HR1c.

5.3.3.4 ^{15}N HR1c + HR1a + HR1b

To investigate whether HR1b can interact with HR1c in the presence of HR1a, HR1b titrations were carried out on the NMR sample containing HR1c and HR1a in a 1:6 ratio. Some small shift changes were observed and were clustered at the first and second contact sites and specifically involved Val228 and Val235 (Figure 5.6). The greatest shift change was that of Leu267 in helix 2, which is consistent with the observed results in the HR1b titrations. None of the other HR1b contact site three residues shifted significantly, suggesting that HR1b binding to HR1c in complex with HR1a may not occur identically to when HR1b binds to free HR1c. The other explanation however could be that HR1a (as dimer) and HR1b, which are both in six-fold excess to HR1c, may preferentially interact with each other, i.e. HR1a interacts with HR1b, preventing it interacting with HR1c. Regardless, the lack of major shift changes suggests that HR1b does not form many further contacts with HR1c in the presence of HR1a. Additionally, reduced intensities on addition of HR1b suggest that a complex may form when all HR1 domains are present in solution.

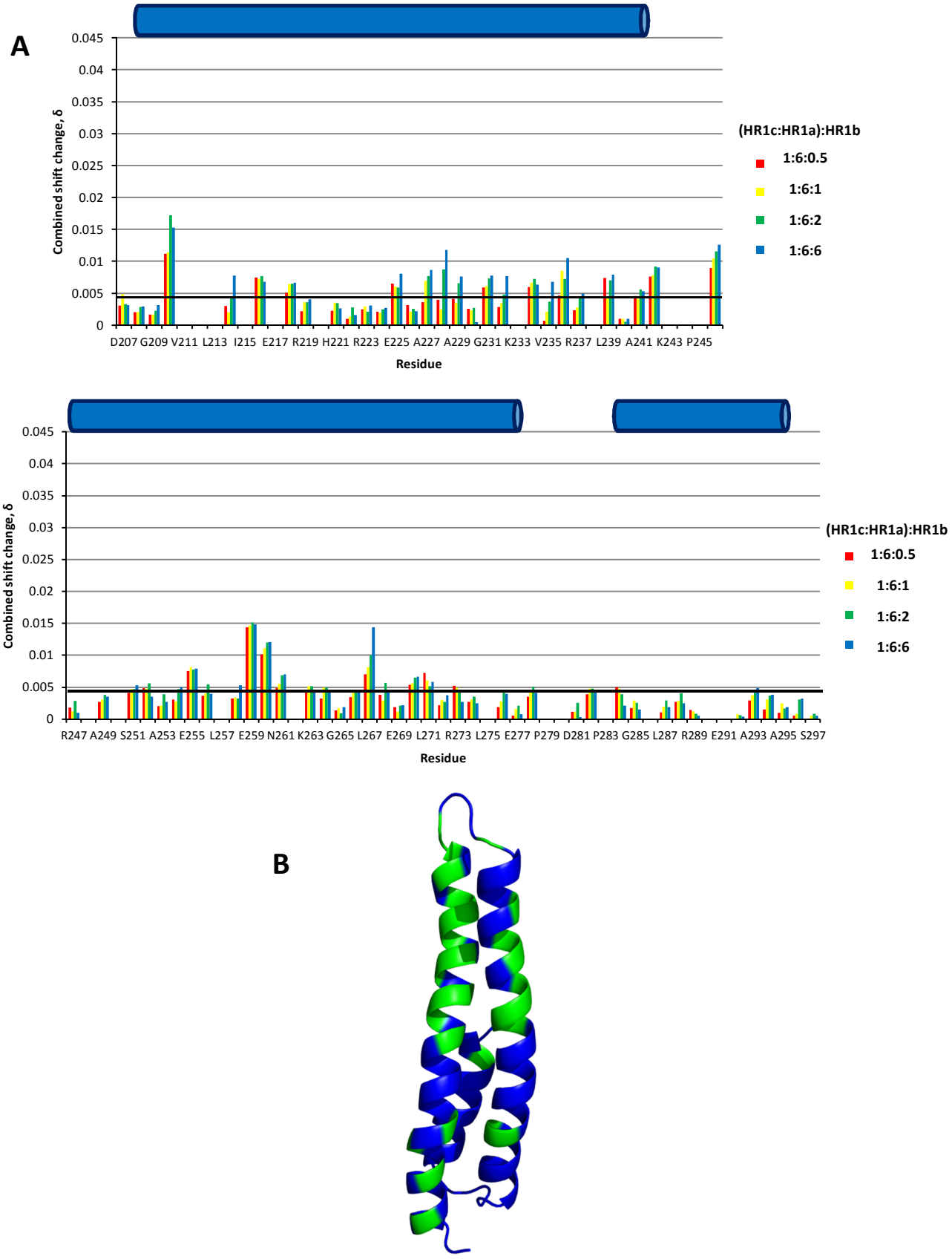


Figure 5.6 Quantitative analysis of ^{15}N HSQC spectra when HR1b was titrated into ^{15}N HR1c + HR1a. (**A**) Bar chart showing the chemical shift changes for each residue of HR1c. The black line represents the mean combined shift change. Mean: 0.0046, standard deviation: 0.004. (**B**) Residues which shifted more than the mean chemical shift change of all residues are coloured green on the structure of HR1c.

5.3.3.5 ^{15}N HR1c + HR1b + HR1a

To investigate whether HR1a can interact with HR1c in the presence of HR1b, HR1a titrations were carried out on the NMR sample containing HR1c and HR1b in a 1:6 ratio. Very small chemical shift changes (~0.01) were observed (Figure 5.7). These changes primarily involved the VEHAVAEGA binding site (but also the VLRLLSAA site), suggesting that HR1a may be displacing HR1b to bind to HR1c. This can also be explained in terms of HR1a and HR1b existing in a dynamic, linked equilibrium with HR1c where more HR1a than HR1b molecules bind to HR1c at a given time (the changes seen here reflect a population average). As before, intensity changes were seen but were not selective.

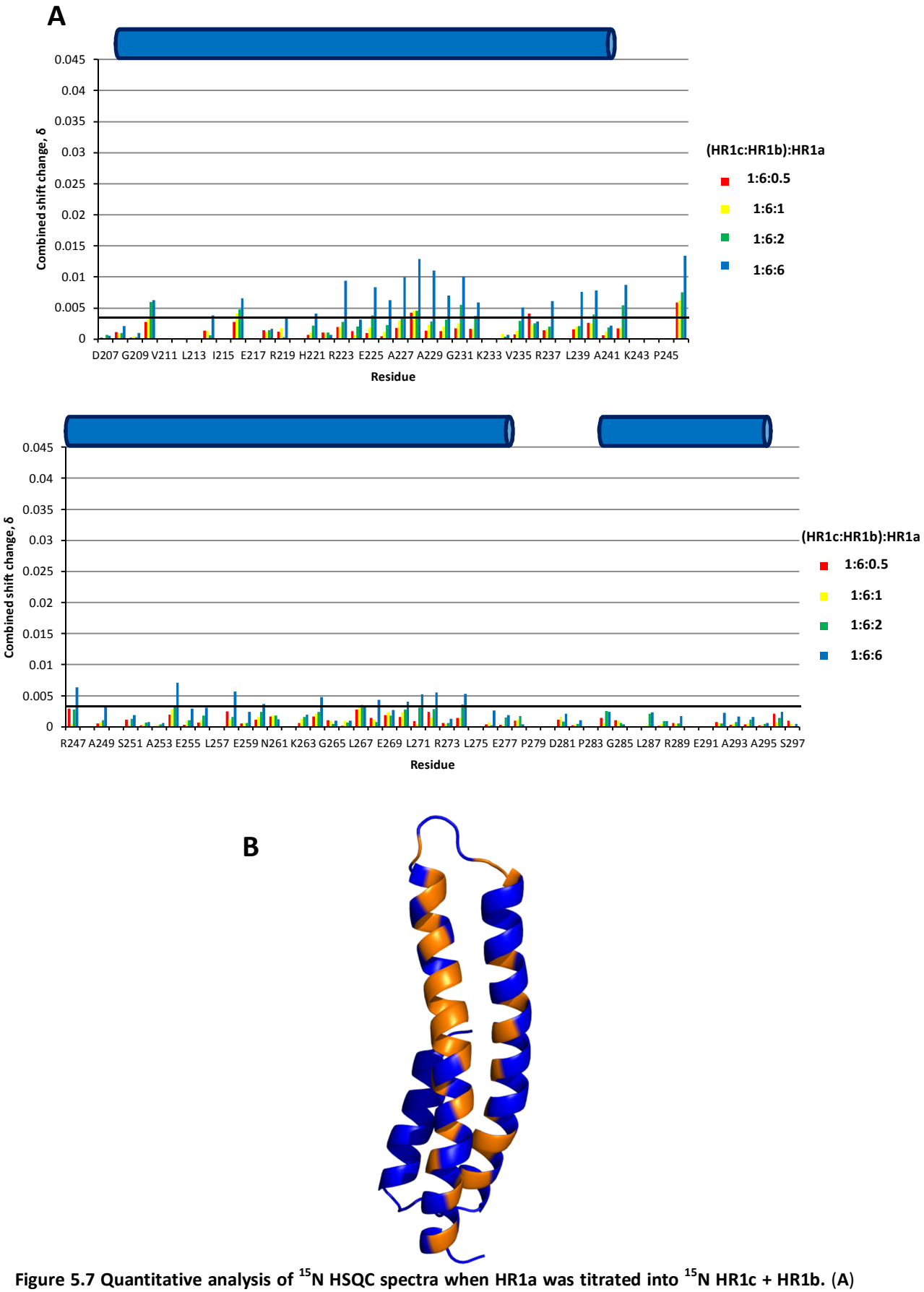


Figure 5.7 Quantitative analysis of ^{15}N HSQC spectra when HR1a was titrated into ^{15}N HR1c + HR1b. (**A**) Bar chart showing the chemical shift changes for each residue of HR1c. The black line represents the mean combined shift change. Mean: 0.0034, standard deviation: 0.0034. (**B**) Residues which shifted more than the mean chemical shift change of all residues are coloured orange on the structure of HR1c.

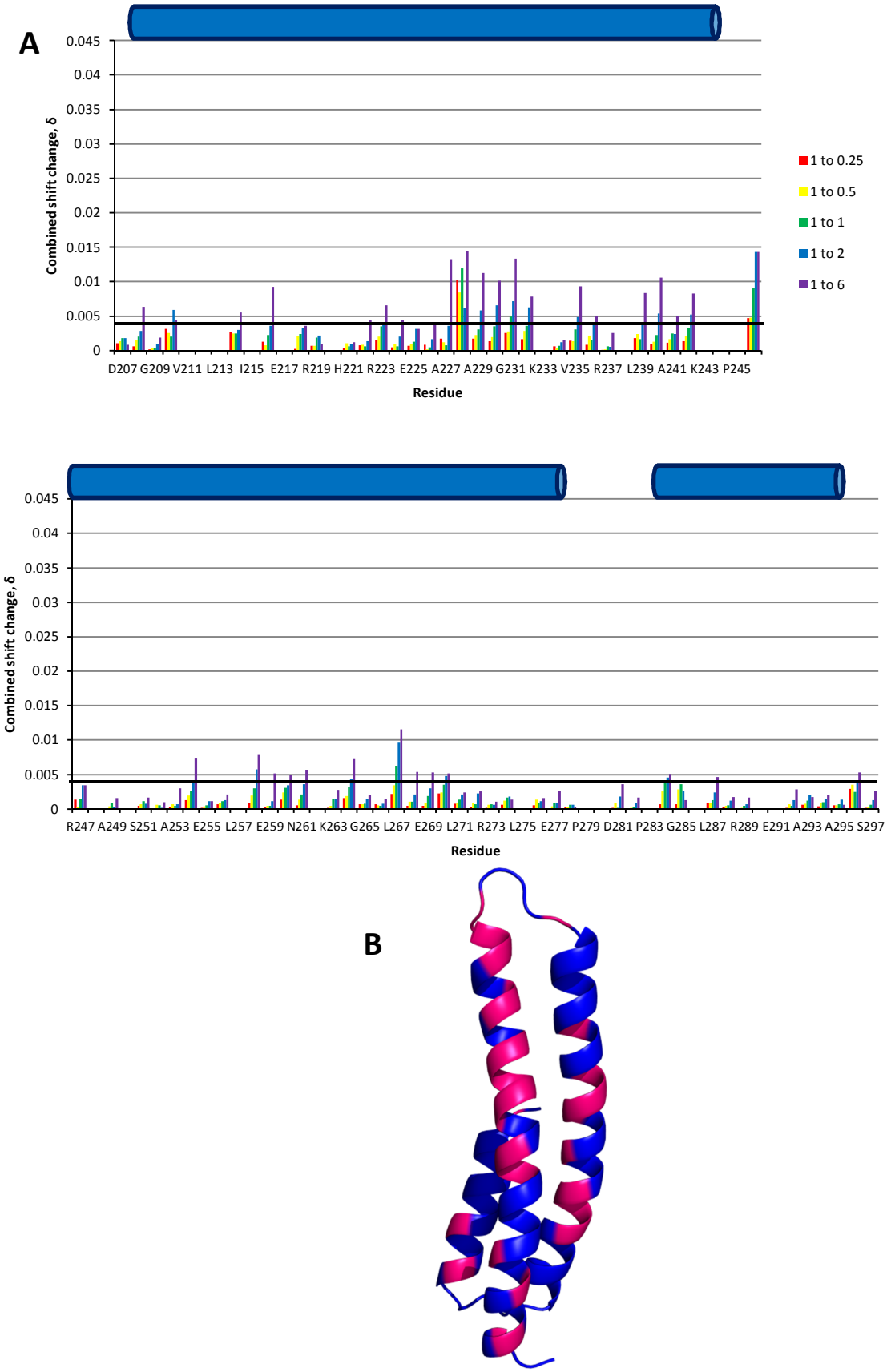
5.3.3.6 ^{15}N HR1c + HR1ab

The interactions of HR1a and HR1b with HR1c were also investigated by titrating the di-domain HR1ab into ^{15}N HR1c. Small shift changes were observed with a magnitude similar to HR1b, suggesting that HR1ab, like HR1b, binds more weakly to HR1c than HR1a does (Figure 5.8). This also suggests that HR1b hinders binding of HR1a to HR1c (possibly because HR1a and HR1b interact with each other). At the titration end-point, the concentration of HR1ab was approximately six times higher than the K_d of HR1ab dimerisation, i.e. a HR1ab was mainly dimeric. This suggests that dimeric HR1ab can interact with HR1c.

The majority of the observed shift changes were clustered in the same two sites seen in the case of HR1a. The first site ranged from Val224 to Ala232 with the greatest shift change being that of Gly231, which is consistent with what was observed previously with the single HR1a domain. Interestingly Val228 did not shift in a single direction but 'moved around', possibly due to distinct interactions with both HR1a and HR1b. The second contact site ranged from Val235 to Ala242, i.e. the end of helix 1 and part of the inter-helical loop. The shift of Ser240 was the greatest however the Leu239 shift was of similar magnitude, which was once again consistent with the HR1a titration results. Some additional shift changes were seen at a third site which matched the other HR1b interacting interface. These shift changes included Thr258, Asn261, Leu264 and Leu267 which were observed previously with HR1b. Interestingly Leu271 did not shift much in these HR1ab titrations in contrast to the HR1b titrations. Regardless, when the HR1a, HR1b and HR1ab chemical shift mapping data are put together, we can conclude that HR1ab can weakly interact with HR1c and that the interaction is driven by HR1a with possible cooperative binding by HR1b.

Line broadening was also observed as indicated by the reduced intensities, suggesting that a complex forms between HR1ab and HR1c. The intensities were reduced by as much as 40-50% in contrast to the smaller percentage changes seen in the experiments above. This suggests that when all three HR1 domains are present, and HR1a and HR1b are tethered by a linker sequence, they can form a complex that is in intermediate exchange (there could also be two ongoing exchange processes; formation of the trimer and HR1a dimerisation). This also highlights the effects of tethering the domains. In cells the three HR1 domains are covalently linked so the effective molarity is much higher than when they exist as individual

domains *in vitro*. This may explain why the HR1abc trimer is the predominant species at 400 µM (according to AUC) even though the NMR titrations suggest a very weak interaction.



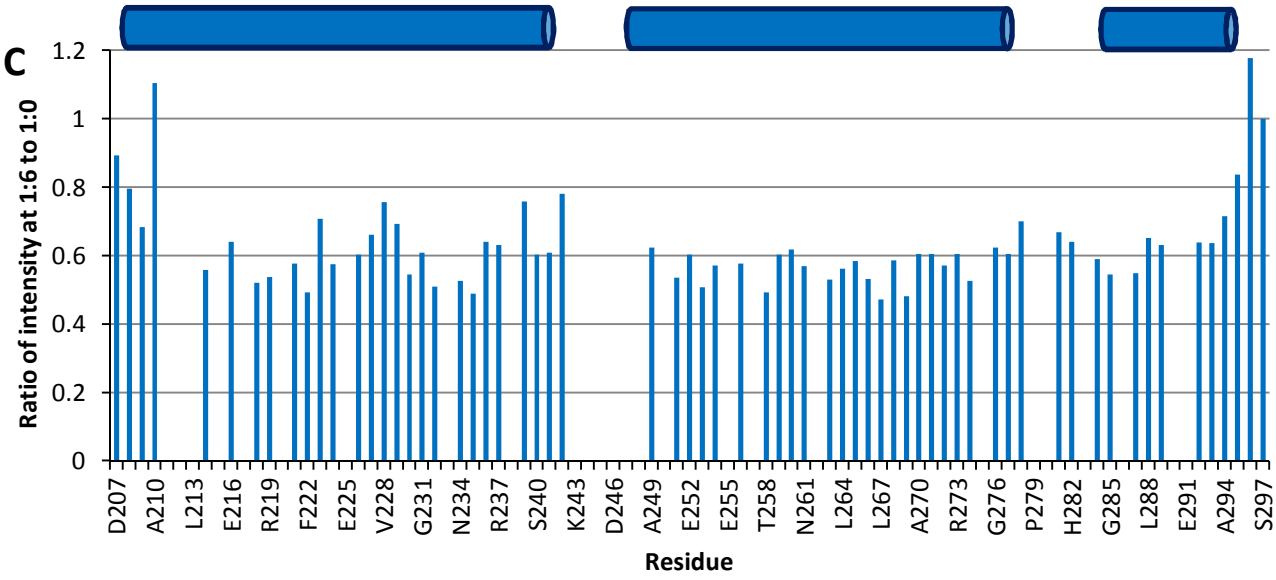


Figure 5.8 Quantitative analysis of ^{15}N HSQC spectra when HR1ab was titrated into ^{15}N HR1c. (**A**) Bar chart showing the chemical shift changes for each residue of HR1c. The black line represents the mean combined shift change. Mean: 0.004, standard deviation: 0.0035. (**B**) Residues which shifted more than the mean chemical shift change of all residues are coloured magenta on the structure of HR1c. (**C**) Ratio of intensity at 1:6 to 1:0.

5.3.4 ^{15}N PRK1 HR1ab

5.3.4.1 Purification of ^{15}N HR1ab

^{15}N-labelled PRK1 HR1ab was expressed in M9 minimal medium and was purified as shown previously for the unlabelled protein. An NMR sample was made up at a concentration of 357 µM.

5.3.4.2 Optimising the ^{15}N HR1ab HSQC

The ^{15}N HR1ab recorded at a concentration of 357 µM suffered from poor spectral dispersion as a result of chemical exchange. Chemical exchange was suspected from the AUC results (Chapter 3), therefore it was decided to dilute the sample and record another HSQC spectrum. New peaks appeared as the concentration was lowered (Figure 5.9). At first glance the ^{15}N HSQC of HR1ab overlays with the ^{15}N HSQC of HR1b (provided by Dr. Helen Mott) (Figure 5.11). Most of the peaks which appeared when diluting HR1ab do not correspond to peaks in the HR1b spectrum, suggesting that those peaks are of HR1a

backbone amides and their linewidth broadening is due to intermediate exchange between the monomeric and dimeric states. Approximately 50 new peaks appeared out of a total of ~120 peaks missing of HR1a (and inter-domain linker) residues when the concentration changed from 357 μM to 15 μM. However, most of the peaks which appeared are very weak in comparison to the peaks which correspond to HR1b backbone amides. This suggests that there is still considerable intermediate exchange between monomeric and dimeric HR1a, even at 15 μM which is ten-fold below the dimerisation K_d of HR1ab. In an attempt to further improve the spectrum, an HSQC was recorded at a higher concentration of NaCl, which would disrupt any salt bridges that could be involved in HR1a dimerisation. However no considerable change was seen (Figure 5.10).

Some peaks which were only observed in HR1ab and must therefore belong to HR1a (or to the 15 amino acid linker between HR1a and HR1b) were strong in the spectra at all the different concentrations. One of these was identified as the side-chain NH of Trp13, which is located at a distinct $^{1}H/^{15}N$ shift. In HR1ab, the HR1b peaks do not undergo line broadening at high concentrations while HR1a peaks do, suggesting that the two domains exist independently in solution. However on closer inspection, some of the HR1b peaks in the HR1ab spectrum show small shift changes relative to their position in the HR1b spectrum (Figure 5.11A). As HR1b is monomeric, the shift changes must be due to an interaction that exists between HR1a and HR1b. Whether HR1b interacts with monomeric or dimeric HR1a cannot be concluded from these data however.

The chemical shift changes were mapped and are shown in Figure 5.11B. One contact site was identified, spanning residues Arg185 to Leu193 which are located in helix 2 of HR1b. These residues are centred around Gln189 which shows the greatest shift change. Additionally, α-helical i to i+3 periodicity is observed between Arg186 and Gln189 which show large shift changes, while Leu193 is also on the same side as these two residues (and it is the last residue of helix 2). These residues are also solvent exposed (>50%), suggesting this is the contact site on HR1b for HR1a. Interestingly, Val128 and Leu131 (start of helix 1) are close to this contact site and also show shift changes, so perhaps they are also part of the HR1a binding site. Figure 5.11C maps these interactions on the structure of HR1b.

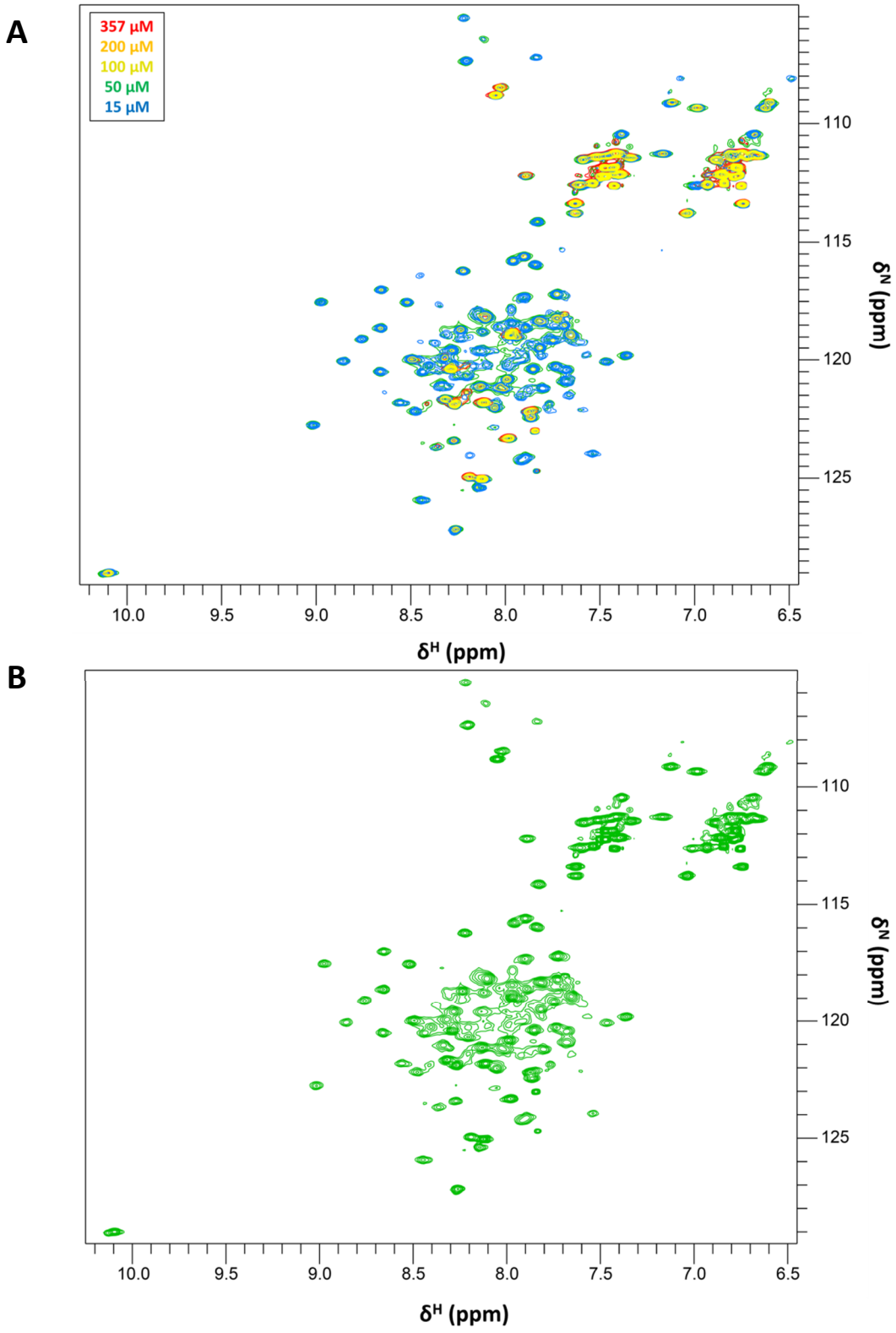


Figure 5.9 ^{15}N-HSQC spectrum of PRK1 HR1ab. (**A**) ^{15}N-HSQC spectra of PRK1 HR1ab in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT and 10% D_2O at concentrations ranging from 357 µM to 15 µM (red to blue) were recorded on an Avance AV800 at 25 ^{o}C. (**B**) ^{15}N-HSQC spectrum of PRK1 HR1ab at 50 µM (concentration used in subsequent studies).

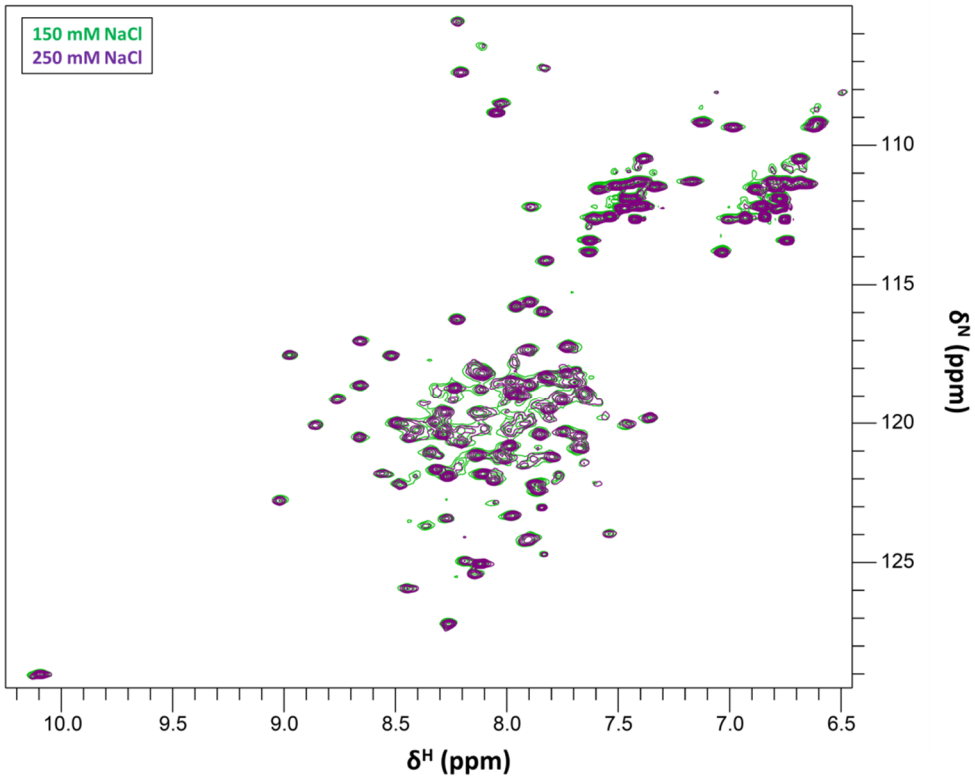


Figure 5.10 Effect of NaCl on quality of the ^{15}N-HSQC spectrum of PRK1 HR1ab. ^{15}N-HSQC spectra of 50 µM PRK1 HR1ab in 50 mM Tris-HCl pH 7.5, 150 mM (green) or 250 mM (purple) NaCl, 5 mM DTT and 10% D_2O were recorded on an Avance AV800 at 25 ^{o}C and are overlaid.

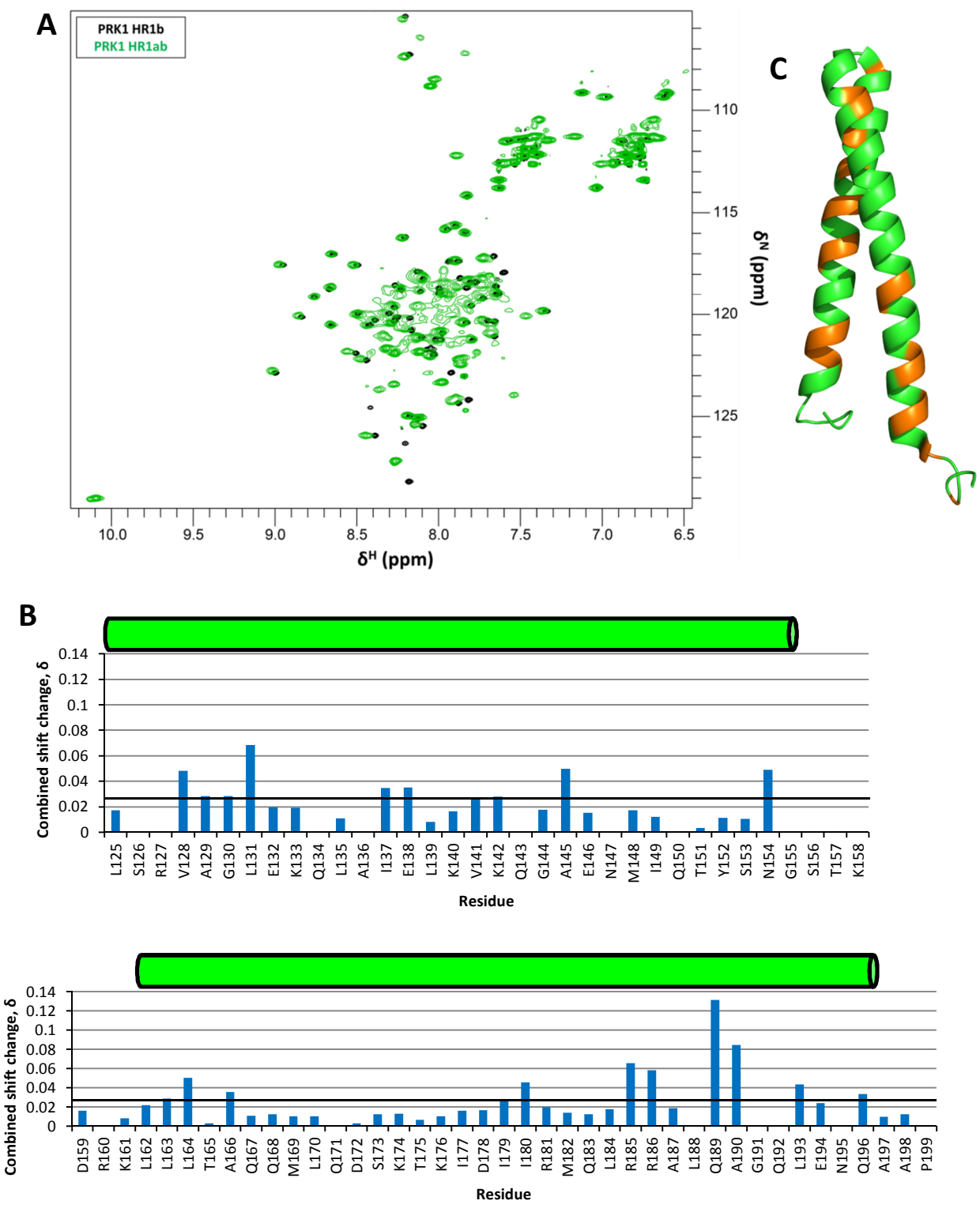


Figure 5.11 Comparison of PRK1 HR1b and PRK1 HR1ab. (**A**) ^{15}N-HSQC spectrum of 50 µM PRK1 HR1ab in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT and 10% D_2O (green) overlaid with PRK1 HR1b spectrum (black) provided by Dr. Helen Mott. (**B**) Bar chart showing the chemical shift changes for each residue of HR1ab in comparison with HR1b. The black line represents the mean combined shift change. Mean: 0.0026, standard deviation: 0.0023. (**C**) Residues which shifted more than the mean chemical shift change of all residues are coloured orange on the structure of HR1b.

5.3.4.3 ^{15}N HR1ab + HR1c

To investigate the binding of HR1c to HR1ab, titrations were carried out using an NMR sample of ^{15}N HR1ab at 50 µM, since at this concentration there is considerable number of new peaks appearing in comparison to higher concentrations. No significant shift changes were observed. A few shift changes were observed in helix 2, primarily those of Leu163, Ala166 and Gln167 but they were very small (Figure 5.12). A few other shift changes were not clustered so cannot be considered important in the binding of HR1c. It is possible that HR1c is interacting with HR1a residues whose backbone amide peaks are not observable in the HSQC (and the results above suggest that HR1a interacts with HR1c more strongly than HR1b does). HR1ab is also mainly monomeric at 50 µM, so it is possible that no shift changes were observed because HR1c prefers to interact with dimeric HR1ab; the reciprocal titration suggested that HR1c binds to dimeric HR1ab and was also characterised by line broadening (suggesting the formation of a larger complex). Unlike the reciprocal ^{15}N HR1c + HR1ab experiment, intensity changes were minimal suggesting that HR1c does not interact with monomeric HR1ab. The significant line-width broadening in HSQC of dimeric HR1ab would in theory prevent mapping the binding interface of HR1c on the di-domain, however this has not been investigated.

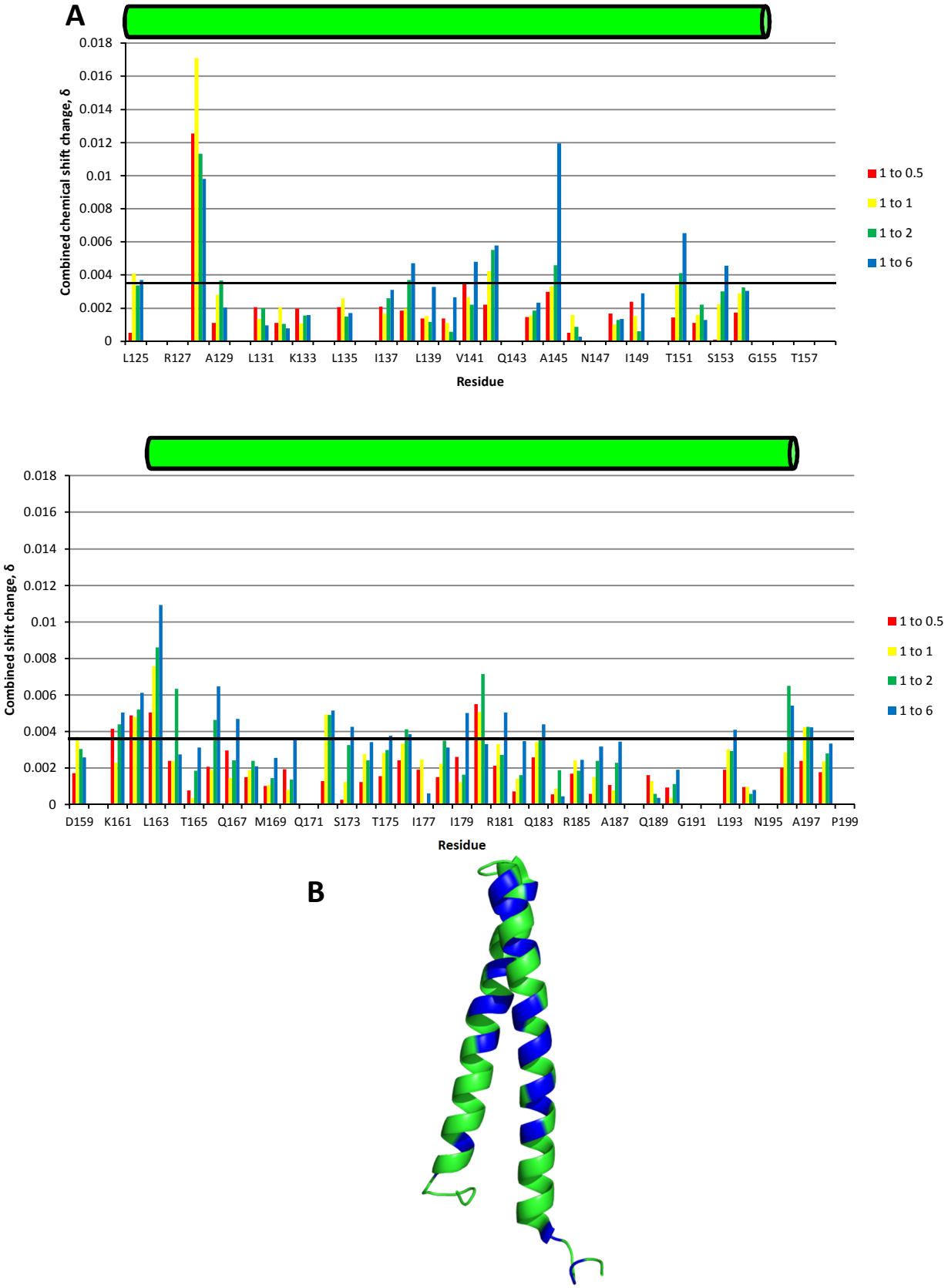


Figure 5.12 Quantitative analysis of ^{15}N HSQC spectra when HR1c was titrated into ^{15}N HR1ab. (**A**) Bar chart showing the chemical shift changes for each HR1b residue in HR1ab. The black line represents the mean combined shift change. Mean: 0.0036, standard deviation: 0.0024. (**B**) Residues which shifted more than the mean chemical shift change of all residues are coloured blue on the structure of HR1b.

5.3.5 Cross-linking experiments

5.3.5.1 Theory

Weak protein-protein interactions pose a hurdle when they are being investigated, as they are reversible and the concentration of purified proteins is often not much higher than the K_d of the interaction, meaning that protein samples are often a mixture of bound and unbound species. In turn, polydispersity hinders detailed structural investigation.

Cross-linking agents are becoming increasingly popular as tools that can be used to lock proteins in a particular conformation or stabilise a weak, transient protein-protein interaction (López-Alonso *et al.*, 2009; Hu *et al.*, 2012; Lepvrier *et al.*, 2014). The aim was therefore to investigate whether cross-linking agents could be used to stabilise monomeric or oligomeric HR1a, HR1ab and HR1abc, which would provide an understanding of the species formed by intra- and inter-molecular interactions and also ideally produce a monodisperse sample that would facilitate structural studies.

5.3.5.2 Chemistry of the different cross-linking agents

1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) is a zero-length, heterobifunctional cross-linker, which can conjugate carboxylates (-COOH) to primary amines (-NH_2) via an amide bond. It is a zero-length cross-linker because it does not become part of the cross-link. As a result, it is considered to have increased cross-linking specificity over other cross-linkers, as any initial modification of a carboxylate that does not successfully lead to a reaction with a primary amine leads to the regeneration of the carboxylate group (Figure 5.13).

Dimethyl suberimidate (DMS) and dimethyl pimelimidate (DMP) are homobifunctional cross-linkers with identical lysine-reactive groups at either end. DMS and DMP only differ in the length of the spacer arm, with DMP being a methylene group shorter. As their spacer arms are 9.2 Å and 11.0 Å long, they theoretically have a higher chance to cross-link proteins non-specifically.

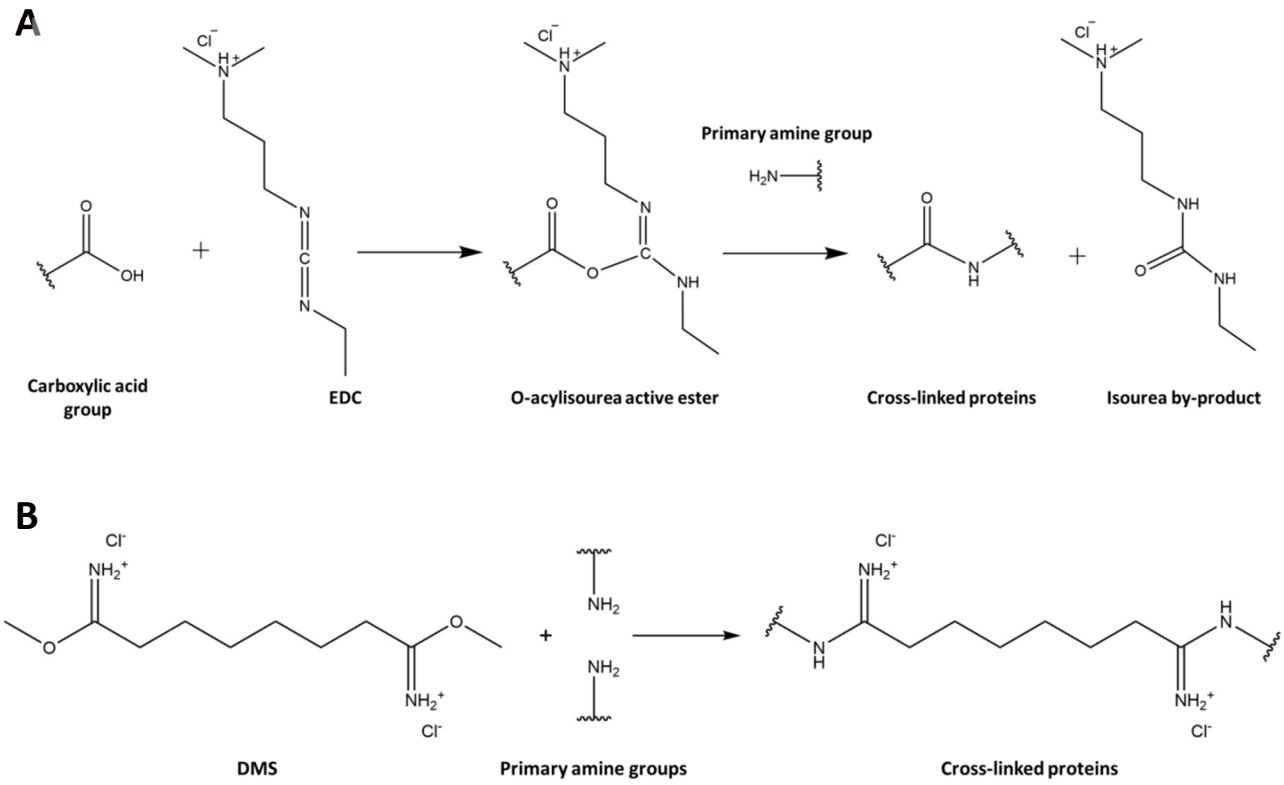


Figure 5.13 Chemistry of EDC and DMS cross-linking agents. The reaction schemes of (**A**) EDC and (**B**) DMS are shown.

5.3.5.3 Results

The first step was to perform small scale trials at different cross-linking agent concentrations. Three concentrations of each cross-linker were chosen and they were tested against HR1a, HR1ab and HR1abc. Following incubation at 30 min or 5 h at room temperature or overnight incubation at 4 ^{o}C, gel samples were taken and were analysed by SDS-PAGE. The concentration of HR1a and HR1ab was 1.5 mM, which is ten-fold above the dimerisation K_d of these proteins. The concentration of HR1abc was 0.7 mM; at this concentration HR1abc is predominantly trimer in solution. These protein concentrations aimed to allow the cross-linking and stabilisation of oligomeric complexes that could be further investigated structurally.

Figure 5.14 shows the HR1a cross-linking trials. At 10 mM, EDC caused the formation of a 25 kDa species which was identified as an HR1a dimer, while a bigger species of ~35-40 kDa was also observed (probably a trimer). The majority of the protein appeared to run as a monomer on the gel, suggesting that either it was unmodified or intramolecularly cross-linked. At 20 mM EDC the concentration of monomer was significantly reduced in

comparison to 10 mM EDC. The levels of the dimer were also reduced while a range of high order oligomeric species formed. At 40 mM EDC, less monomer, less dimer and a greater proportion of higher order oligomers formed. Since HR1a is known to only exist as monomer and dimer in solution, the oligomers seen here are non-physiological, suggesting that EDC is inducing non-specific cross-linking. The dimer seen on the gels may therefore not be a stabilised version of the dimer which forms in solution. Obtaining a monodisperse sample of the dimer which can be used in biophysical analysis is therefore unobtainable by this method.

In the case of DMS and DMP, monomer and dimer were the main species observed across all cross-linking agent concentrations (Figures 5.14B-C). While there was no major difference between DMS and DMP reactions (as indicated by dimer levels on the gel), most

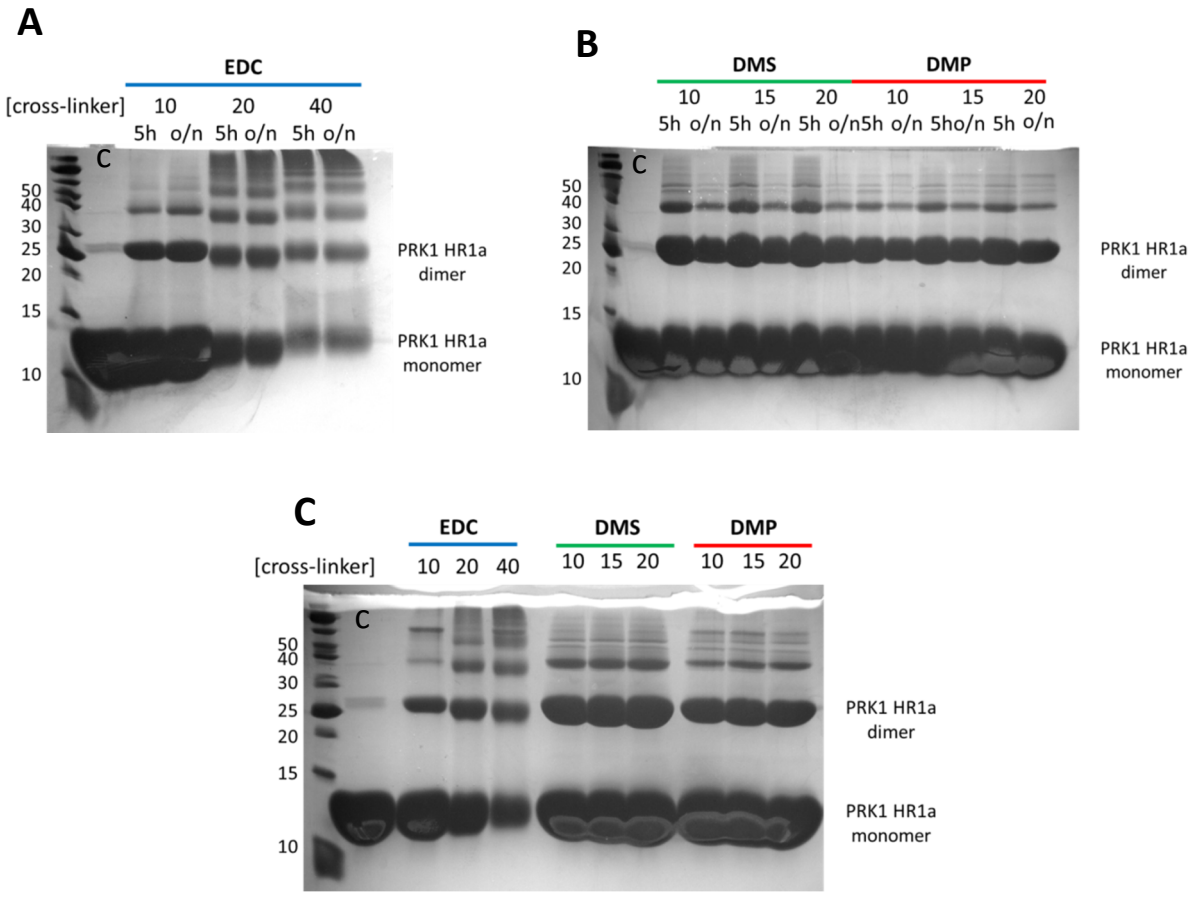


Figure 5.14 PRK1 HR1a cross-linking trials. Cross-linking was performed at a range of EDC, DMS or DMP concentrations. (**A**) EDC reactions after 5 h at room temperature or overnight at 4 ^{o}C. (**B**) DMS and DMP reactions after 5 h at room temperature or overnight at 4 ^{o}C. (**C**) EDC, DMS and DMP reactions after 30 min. C stands for control.

of the protein remained monomeric, either because it didn't react with the cross-linking agent, or because it was stabilised intra-molecularly. Higher order oligomers formed at low levels and were generally lower when DMP was used, suggesting that the 'longer arm' DMS causes slightly higher levels of non-specific cross-linking.

Cross-linking of PRK1 HR1ab with EDC produced similar results to HR1a cross-linking with EDC. Some dimer was observed at 10 mM EDC although most of the protein remained monomeric (Figures 5.15A and 5.15C). At higher concentrations of EDC, the levels of monomer and dimer decreased and higher molecular weight species formed, once again, suggesting non-specific cross-linking took place.

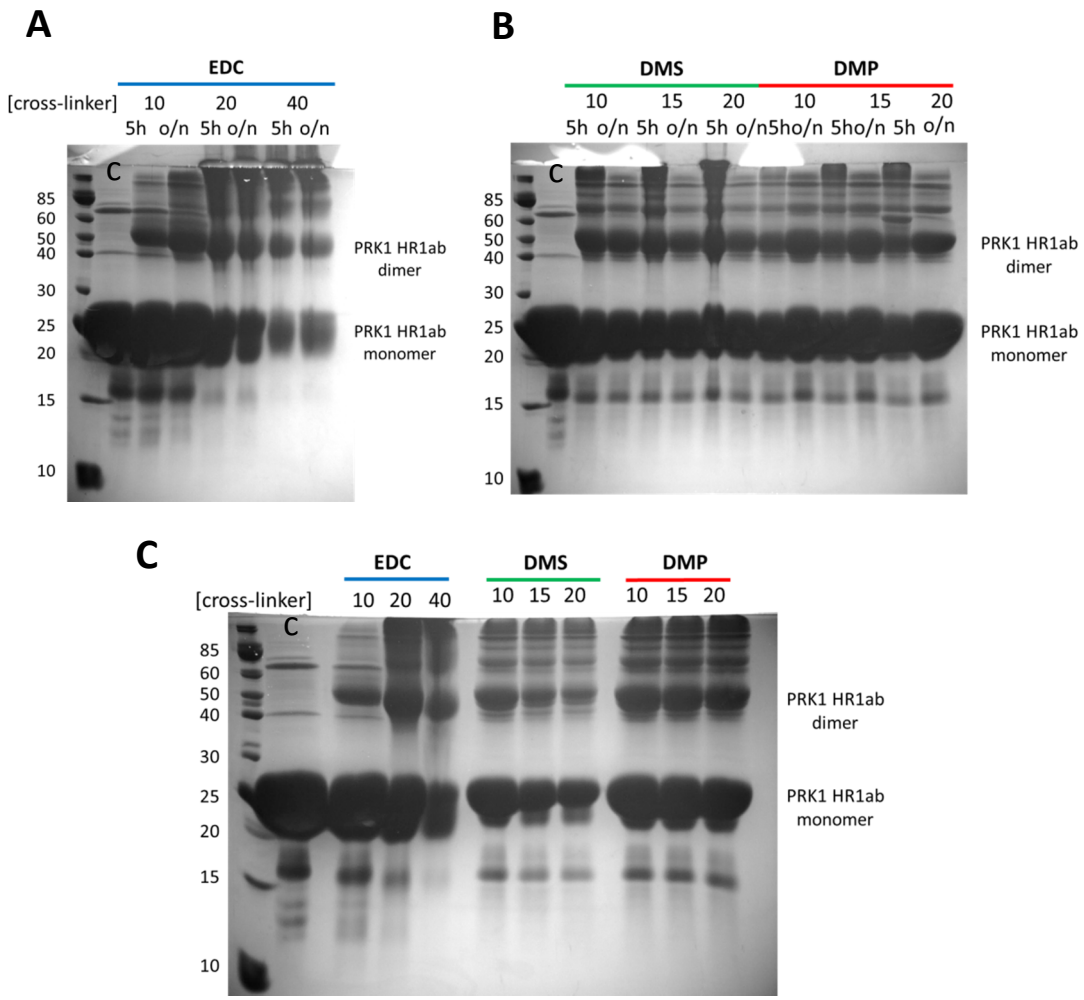


Figure 5.15 PRK1 HR1ab cross-linking trials. Cross-linking was performed at a range of EDC, DMS or DMP concentrations. (**A**) EDC reactions after 5 h at room temperature or overnight at 4 °C. (**B**) DMS and DMP reactions after 5 h at room temperature or overnight at 4 °C. (**C**) EDC, DMS and DMP reactions after 30 min.

Cross-linking of PRK1 HR1abc with EDC resulted in mainly monomer and low levels of dimer at 10 mM EDC, while at higher EDC concentrations HR1abc formed higher-order oligomers which ran as a smear on the gel or were stuck in the stacking gel wells (Figures 5.16A and 5.16C). With DMS and more so with DMP, the major species was the monomer so it was possible that this was a stabilised HR1abc monomer (Figures 5.16B and 5.16C).

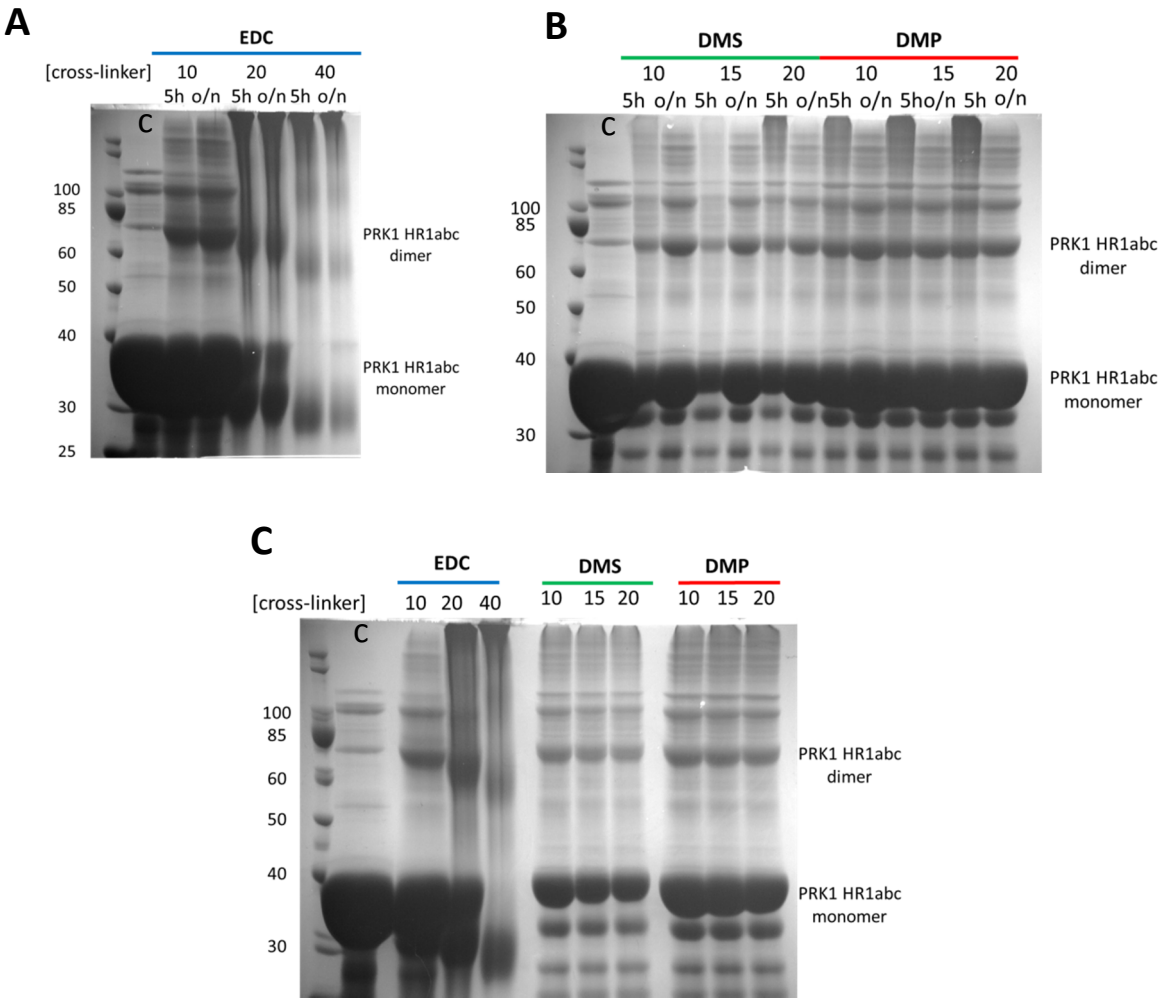


Figure 5.16 PRK1 HR1abc cross-linking trials. Cross-linking was performed at a range of EDC, DMS or DMP concentrations. (**A**) EDC reactions after 5 h at room temperature or overnight at 4 ^{o}C. (**B**) DMS and DMP reactions after 5 h at room temperature or overnight at 4 ^{o}C. (**C**) EDC, DMS and DMP reactions after 30 min. C stands for control.

The stabilisation of HR1a, HR1ab and HR1abc monomers by these cross-linkers was investigated by setting up cross-linking reactions of the best conditions identified above and the products analysed by size exclusion chromatography in comparison to a control reaction. The elution volume of the monomer remained identical between control and cross-linked samples suggesting that monomers seen in the trials above were not stabilised.

In order to stabilise the interaction between HR1ab and HR1c, cross-linking studies were set up (Figure 5.17). While a low level of a species corresponding to the molecular weight of HR1abc can be seen in both EDC and DMS reactions (particularly in the EDC reaction), HR1c itself is cross-linked by both EDC and DMS.

The failure of cross-linking reactions to stabilise oligomers doesn't mean that these oligomers don't form in solution. It simply implies that the appropriate functional groups are not used in the interacting interface or that they are not located at the appropriate positions.

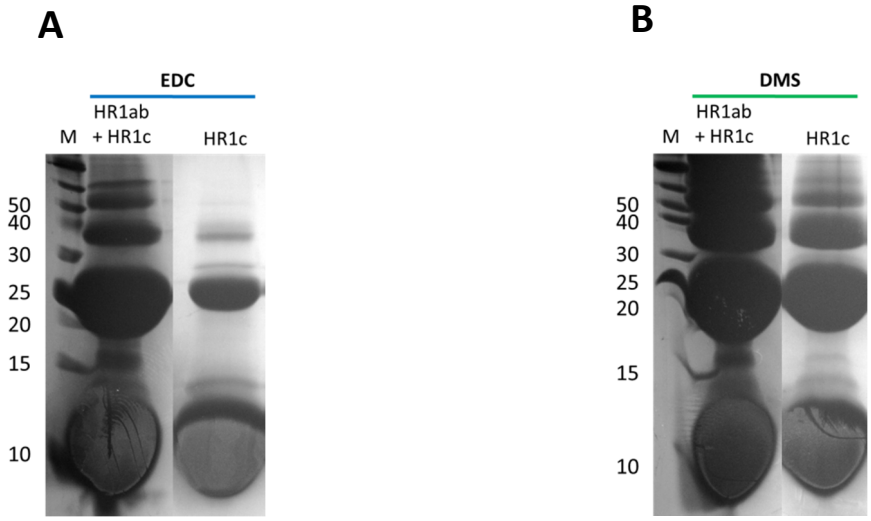


Figure 5.17 Cross-linking trials of PRK1 HR1ab and PRK1 HR1c. Cross-linking reactions were set up with 1.5 mM PRK1 HR1ab mixed with 1.5 mM PRK1 HR1c or with PRK1 HR1c alone (control reaction). (**A**) EDC reactions were performed with 10 mM EDC overnight at 4 ^{o}C. (**B**) DMS were performed with 20 mM DMS overnight at 4 ^{o}C.

5.4 Conclusions

The interactions between the three PRK1 HR1 domains are weak interactions. Therefore a question arises as to whether these weak, transient interactions are truly physiologically relevant. However these weak affinities are of individual domains which are independent in solution. Physiologically, these domains exist in the context of the HR1abc tri-domain and they are covalently linked to each other so that their effective molarity is much higher. Therefore the affinity of these domains for each other *in vivo* is likely to be higher than experiments with individual domains suggest. Furthermore, evidence has been presented here that all three of the PRK1 HR1 domains interact with each other; HR1a with HR1b, HR1a with HR1c and HR1b with HR1c. These interactions may be intra-molecular or inter-molecular. Therefore the effective molarity is expected to be further enhanced as more of these intra- and inter-molecular interactions occur, i.e. there probably is a cooperative effect in the formation of the HR1abc trimer. The HR1abc trimer itself is probably driven by the HR1a·HR1c and HR1b·HR1c interactions, as HR1ab only forms monomer and dimer, and it is the presence of HR1c (which is monomeric by itself) in the context of HR1abc which leads to trimer formation. An important piece of the puzzle that is missing however is the stoichiometry of these interactions; for example, two HR1a domains may interact with a single HR1c domain via each of the two main binding sites on helix 1 or two HR1b domains may interact with HR1c, each on one of two opposite surfaces.

The interaction of HR1a with HR1b may also compete with the interaction of these two domains with HR1c. This is suggested by the inability of the HR1ab to interact with HR1c while monomeric HR1a may interact (early titration points; see section 5.3.3.2). Interestingly none of these contact sites involve residues in helix 3 of HR1c. The function of this unique C-terminal helix therefore remains unknown.

These HR1-HR1 interactions may be weak and transient, however equivalent interactions are important in cellular signalling. Their purpose often involves the rapid activation and ensuing inactivation of signalling pathways through the formation of reversible protein-protein interactions. In this case they could allow HR1abc to transiently exchange between different oligomeric states, each of which may have a particular functional role. The

domains are also covalently linked in full length PRK1 and therefore their effective molarity is much higher. This suggests that these interactions may indeed be physiologically relevant. Other examples of signalling proteins with intramolecular interactions governed by effective molarity include the non-receptor tyrosine kinase src (LaFevre-Bernt, *et al.*, 1998).

In addition to the interactions between different HR1 domains, HR1a has also been shown to dimerise (Chapter 3). The NMR titration data presented in this Chapter suggest that dimeric HR1a and HR1ab can weakly interact with HR1c. Assuming that the HR1abc dimer is formed by HR1a dimerisation and that HR1abc oligomerisation follows a linear pathway from monomer to dimer to trimer, then it is possible that HR1a dimerisation may have a role in bringing two PRK1 molecules in proximity to form the dimer, which is then poised to interact with an additional protomer, for example via HR1c. The dimerisation itself appears to be unaffected by the presence of the other HR1 domains as indicated by the ^{15}N HR1ab/HR1c titration where the HR1a peaks still remain significantly broadened. This is also suggested by the ability of the HR1a dimer to bind to HR1c as suggested by the ^{15}N HR1c/HR1a titration.

A further unknown concerns the oligomerisation of HR1abc which stops at trimer. If the HR1 domain interacting interfaces were not all satisfied in the trimer, then it would be expected that the trimer would form larger oligomers. This doesn't happen however, suggesting that all these contact sites are used in the HR1abc trimer so an additional protomer cannot interact to form a larger oligomer. As discussed in Chapter 3, this may also be indicative of a relatively globular species and not one which linearly oligomerises. An additional aspect which remains unclear is whether HR1abc can have several conformations, even as a monomer, depending on which of the contacts between the HR1 domains form. These different conformations may have several implications including the activation of the PRKs; stabilising HR1a in a conformation that keeps its autoinhibitory sequence away from the PRK1 active site; the preferential binding of a GTPase to a particular HR1 domain.

It is known that binding of RhoA (and Rac1) to PRK1 activates it. These intra- and inter-molecular interactions may therefore serve to keep the HR1abc regulatory region of PRK1 in a 'closed', autoinhibited state. Binding of GTPases, which occur with nM to low µM affinities, may therefore 'open up' the structure, disrupting these HR1-HR1 interactions and

in turn PRK1 oligomers. The effect of GTPase binding on PRK1 oligomerisation is investigated in Chapter 6.

An interesting observation can also be made when looking at the aligned sequence of the three PRK1 HR1 domains (Figure 5.18). Many of the residues in HR1c that interact with HR1a and HR1b correspond to residues in HR1a and HR1b that interact with RhoA and Rac1, respectively. These residues are Glu225^{HR1c}, Val228^{HR1c}, Gly231^{HR1c} and Ala232^{HR1c}, which are conserved as Glu49^{HR1a}/Glu138^{HR1b}, Leu52^{HR1a}/Val141^{HR1b}, Gly55^{HR1a}/Gly144^{HR1b} and Ala56^{HR1a}/Ala145^{HR1b} in HR1a and HR1b, respectively. Interestingly, they are three of only eight residues fully conserved across all three PRK1 HR1 domains, and three out of four residues which are both fully conserved and also used to bind to GTPases. The above suggests that binding of a GTPase to HR1c may be prevented if HR1a and HR1b compete for the same binding site. Conversely, it is possible that binding of a GTPase to HR1c may displace HR1a and HR1b, causing further conformational changes in the HR1abc tri-domain, with potential functional implications. Even if HR1c has no cognate GTPase binding partner, it is intriguing that this same site is used in HR1 domains to bind either to GTPases or to other HR1 domains. This may indicate that this region is thermodynamically less rigid and therefore can accommodate the required local conformational changes that are required for binding.

```
PRK1_HR1a MASDAVQSEPRSWSLLEQLGLAGADLAAPGVQQQLELERERLRREIRKELKLKEGAENLR  60
PRK1_HR1b ---------------------------------ATNLSRVAGLEKQLAIELKVKQGAENMI  149
PRK1_HR1c ------------------------DTQGSPDLGAVELRIEELRHHFRVEHAVAEGAKNVL  236
                                              .   *.:.:  *  : :**:*:

PRK1_HR1a RATTDL----GRSLGPVELLLRGSSRRLDLLHQQLQELHAHVVLPDPAAT----------- 106
PRK1_HR1b QTYSNGSTKDRKLLLTAQQMLQDSKTKIDIIRMQLRRALQAGQLENQAAP----------- 199
PRK1_HR1c RLLSAAKAPDRKAVSEAQEKLTESNQKLGLLREALERRLGELPADHPKGRLLREELAAASS 297
          :  :       : :  .:  *  *. ::.:::  *..        .  .
```

Figure 5.18 Sequence alignment of PRK1 HR1 domains. Asterisks highlight identical residues at all positions, colons highlight conservation of a strong group and periods highlight conservation of a weak group. Interacting residues in the RhoA-PRK1 HR1a (Maesaki *et al.*, 1999) and Rac1-PRK1 HR1b (Modha *et al.*, 2008) structures are highlighted in yellow and green, respectively.

Chapter 6 GTPase interactions with HR1 domains

6.1 Introduction

6.1.1 Known interactions

PRK1 was first identified as an effector of RhoA (Watanabe *et al.*, 1996) and Rac1 (Vincent & Settleman, 1997), but while the N-terminal region of PRK1 was shown to be mediating these interactions, the exact HR1 domains binding to these GTPases remained unknown until work by Flynn *et al.* (1998) and Owen *et al.* (2003). Maesaki *et al.* (1999) provided the first structural characterisation of an HR1 domain, that of HR1a in complex with RhoA. The NMR solution structure of Rac1 in complex with HR1b (Modha *et al.*, 2008) and mutational analysis of the RhoA/HR1a interface by Hutchinson *et al.* (2011) provided much needed insight into the contact site used in these interactions. A further study by Hutchinson *et al.* (2013) established that the PRKs are effectors of the entire Rho subfamily (RhoA, RhoB and RhoC), while more recently a mass spectrometry study from mouse tissues has suggested that RhoD may also interact with PRK1 (Paul *et al.*, 2017).

Table 6.1 summarises the published affinities of PRK1 HR1 domains for Rho GTPases. It is worth noting that HR1 domains of TOCA family proteins interact with Cdc42 and TC10 with 10-100-fold lower affinities (~10 μM) than Rho GTPases, suggesting that features other than the overall domain fold are crucial for binding.

GTPase	HR1 domain	Affinity (nM)
RhoA	PRK1 HR1a	150[1], 150[3]
	PRK1 HR1b	>1000[2], 1800[3]
	PRK1 HR1ab	140[1]
RhoB	PRK1 HR1a	52[1]
	PRK1 HR1b	-
	PRK1 HR1ab	15[1]
RhoC	PRK1 HR1a	82[1]
	PRK1 HR1b	-
	PRK1 HR1ab	90[1]
Rac1	PRK1 HR1a	169[2]
	PRK1 HR1b	68[2]
	PRK1 HR1ab	65[2]

Table 6.1 Summary of published G protein/PRK1 HR1 domain affinity data. The K_d values of the individual HR1a or HR1b domains, as well as the HR1ab di-domain are listed. Hyphens represent binding data that are not available. The number in superscript denotes the paper from which the data were obtained. 1 – Hutchinson *et al.*, 2011; 2 – Owen *et al.*, 2003; 3 – Blumenstein & Ahmadian, 2004.

6.1.2 Aims

The possibility that there are other GTPases which bind to the HR1 domains but remain unknown is very likely. Predicting which GTPases may interact with the PRK HR1 domains, as well as how strongly they may interact, is not a straightforward task. This is because binding of HR1 domains appears to not only depend on sequence specificity, but rather also on their structural stability and rigidity (Hutchinson *et al.*, 2013). Currently, only members of the Rho and Rac subfamilies of the Rho family of GTPases are known to bind to PRK effector domains, while the diversity of the HR1 domains extends to the Cdc42 subfamily when the TOCA, CIP4 and FBP17 proteins are also considered. It is possible that members of other subfamilies can interact with the PRK HR1 domains. Such interactions may remain unknown because these GTPases and the HR1 domains may interact very weakly and also because the other GTPase subfamilies are less well-studied. A recent study employed mass spectrometry

to identify effectors of six well-studied Rho GTPases, RhoA, RhoB, RhoC, Rac1, Cdc42 and RhoD (Paul *et al.*, 2017). Interestingly, PRK1 was pulled down not only by some of its known interacting GTPases, but also with RhoD. While the specific HR1 domain interacting with RhoD was not the focus of this mass spectrometry study, these results highlight that our knowledge of GTPase/effector interactions remains nebulous.

This chapter aims to identify interacting partners of the HR1c domain. In addition, it aims to provide preliminary characterisation of GTPase interactions with the HR1abc tri-domain, as well as a measure of the affinity of RhoA for the HR1abc tri-domain in an attempt to understand whether interactions between the HR1 domains affect binding of GTPases. Furthermore, this Chapter aims to investigate if the HR1a dimer can interact with RhoA or not. The above investigations aim to provide further understanding of the HR1abc tri-domain structure and attempt to link GTPase interactions to the discovery of HR1 oligomerisation presented in Chapter 3.

6.2 Investigating new interaction partners between the Rho family and the HR1 domains

6.2.1 Choice of GTPases to test

To test binding to GTPases pull-downs were conducted between GST-tagged GTPases and untagged HR1 effectors. In the interest of time, binding of all 23 members of the Rho family to the HR1c domain could not be investigated. RhoA was chosen as a positive control when optimising the pull-downs as it can interact with HR1a. It also had an additional role as a negative control for HR1c pull-downs, as RhoA has been shown not to interact with PRK1 HR1c (Flynn *et al.*, 1998). RhoD was also chosen, a decision primarily made based on its recent identification as an interacting partner of PRK1 (Paul *et al.*, 2017), but also due to its localisation in the endosomes of cells, where PRK1 is also known to localise. Until recently, PRKs were thought to be exclusive effectors of the Rho and Rac subfamily GTPases. RhoD is

not a member of these subfamilies so it was interesting to see if PRK interactions extend to other members of the Rho family.

Additionally, the putative interaction of Cdc42 with PRK3 HR1c, firstly seen in the Part III Master's project (Sophocleous, 2014), was further investigated with the aim of obtaining a K_d of this interaction. Cdc42 also interacts with the HR1 domains of TOCA, CIP4 and FBP17 so a K_d would provide clues as to whether this interaction is of nM affinity like those of RhoA/HR1a and Rac1/HR1b or μM affinity like those of the TOCA family HR1 domains for Cdc42.

6.2.2 Choosing the best expression conditions

Before performing the pull-downs, the expression and solubility of GST-RhoA and GST-RhoD GTPases was assessed with small scale expression (Figure 6.1). Three *E. coli* strains were chosen. These were BL21, Rosetta2 which provides rare codon tRNAs for five amino acids and C41, which carries uncharacterised mutations that help with the solubility of 'toxic' proteins. Both RhoA and RhoD showed increased solubility at 20 ^{o}C than at 37 ^{o}C. RhoA was soluble in all three strains but there was elevated expression in Rosetta2, which was the strain chosen for expression. RhoD was completely insoluble in Rosetta2 and mostly insoluble in BL21 as well. However some soluble protein was seen in the C41 strain which was therefore the strain chosen for expression.

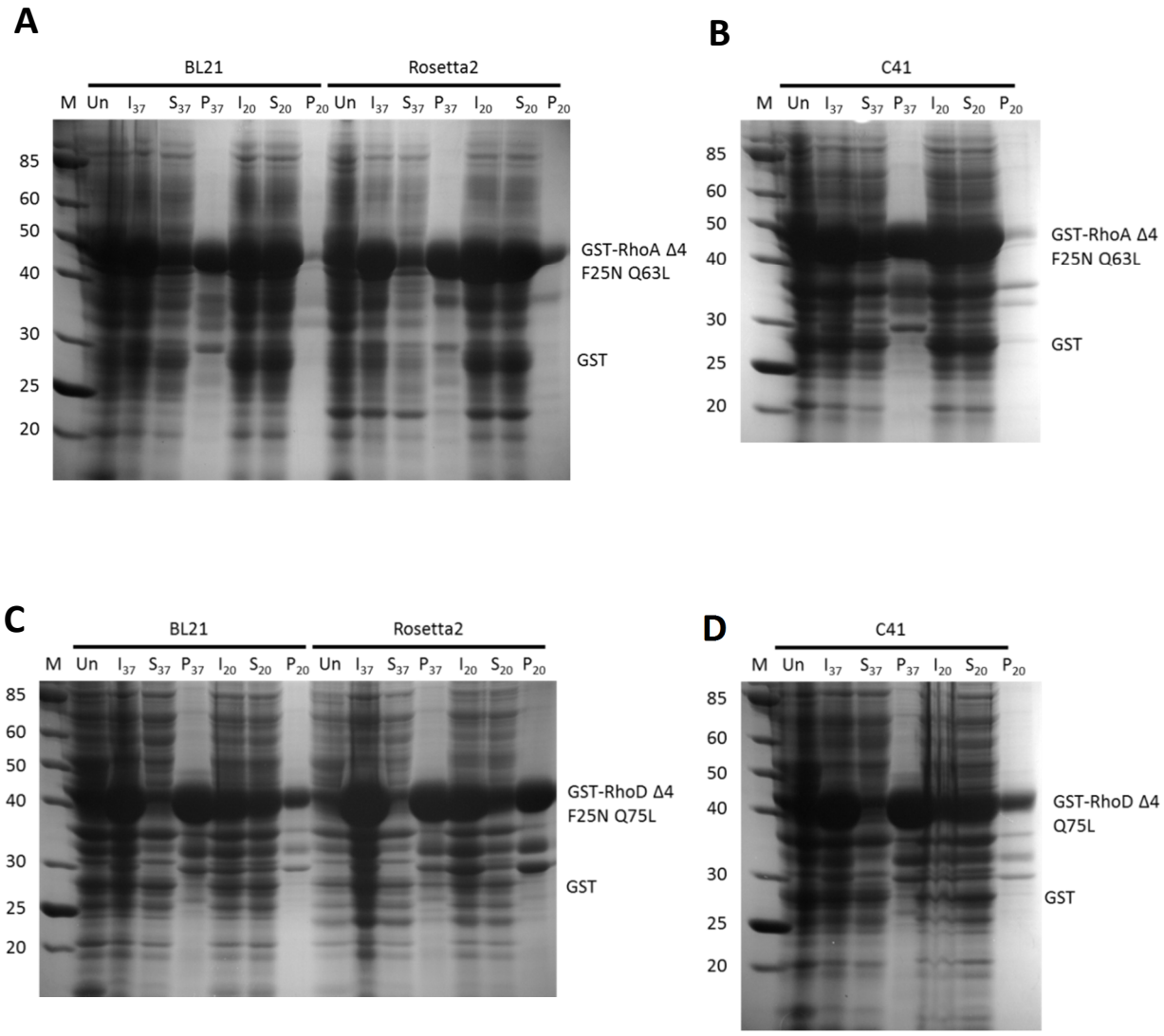


Figure 6.1 SDS-PAGE Analysis of Rho GTPase expression trials. The fusion proteins were expressed in BL21, Rosetta2 and C41 *E.coli* strains. Solubility of (**A**) GST-RhoA Δ4 F25N Q63L and (**B**) GST-RhoD Q75L was assessed by identifying a band of appropriate molecular weight in soluble or pellet fractions. Un for Uninduced sample, In for Induced (Total), S for Induced (Soluble) and P for Induced (Pellet). 20 and 37 denote the temperature at which the expression was carried out at. Laemmli gels were 12%.

6.2.3 Pull-downs

To investigate the binding of HR1 domains to RhoA and RhoD, 2-way pull-downs were performed between N-terminal GST-tagged RhoA and RhoD and untagged PRK1 HR1a, HR1b, HR1c and HR1ab effector proteins. Constitutively active variants of the GTPases were used to allow GTP-loaded GTPases to be expressed and purified from *E.coli* and immediately used in pull-down, without the need for a prior nucleotide exchange step.

RhoA pull-downs (Figure 6.2) indicated that HR1a and HR1ab interact with RhoA as expected. HR1c did not interact with RhoA, which is consistent with earlier observations (Flynn *et al.*, 1998). PRK1 HR1b stains very weakly with Coomassie blue and also interacts with RhoA more weakly than HR1a does, so the HR1b pulled down by GST-RhoA can only faintly be seen.

RhoD was also confirmed to interact with PRK1 (Figure 6.3). Specifically, the RhoD pull-downs showed that HR1a and HR1ab interact with RhoD, i.e. HR1a is the specific effector domain which interacts with this GTPase. RhoD also appears not to interact with HR1b or HR1c, but the result is inconclusive as a result of a limited amount of GST-RhoD immobilised on glutathione agarose beads.

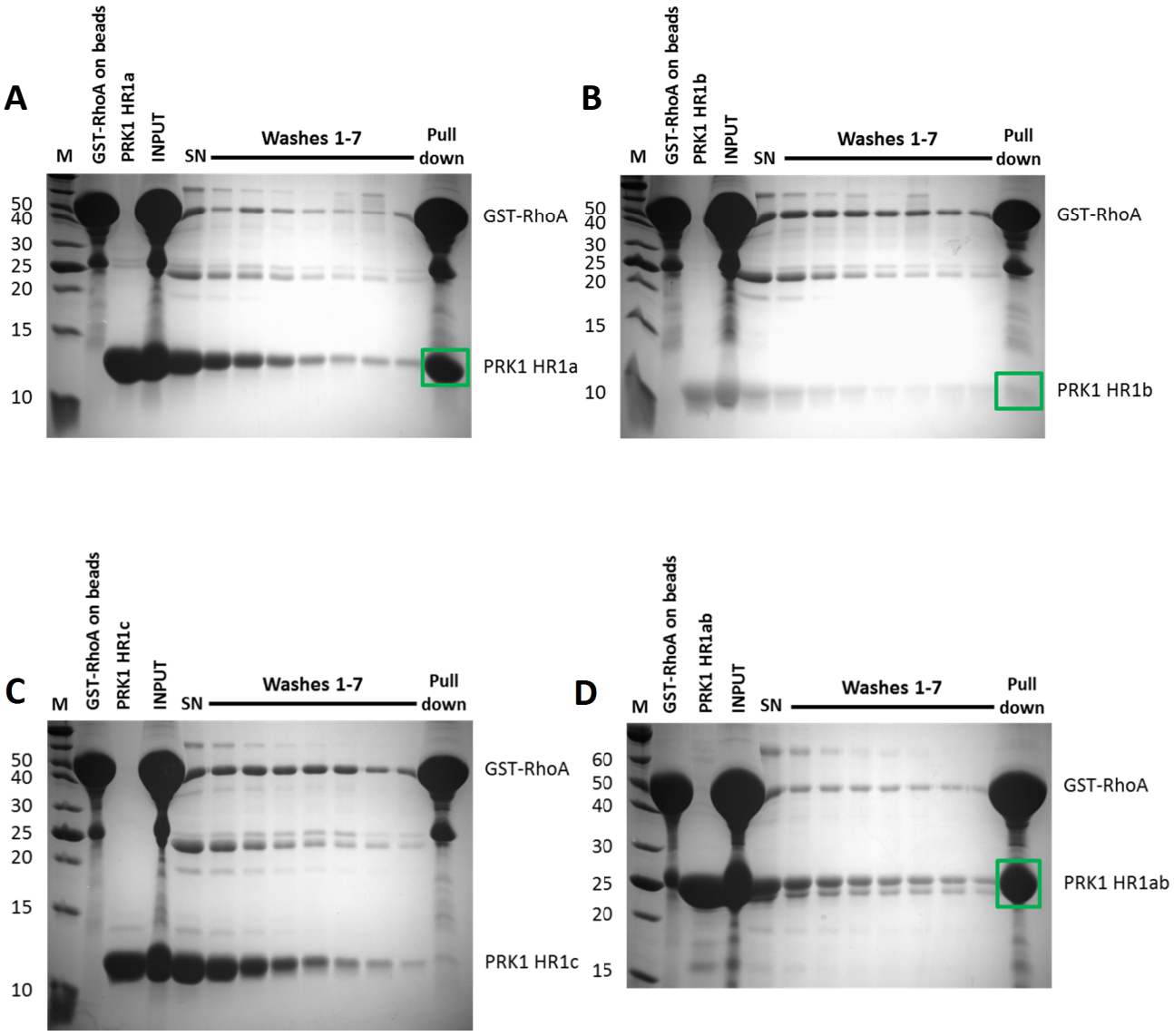


Figure 6.2 RhoA interacts with PRK1. 2-way pull-downs between GST-RhoA Δ4 F25N Q63L and (**A**) PRK1 HR1abc, (**B**) PRK1 HR1b, (**C**) PRK1 HR1c or (**D**) PRK1 HR1ab. HR1 effector protein was added to GST-RhoA immobilised on glutathione agarose beads. Following incubation, unbound protein was washed off. Binding was assessed by SDS-PAGE on gels of the appropriate percentage. M for molecular weight markers and SN for supernatant after incubation. The pull-down sample represents the proteins bound to the glutathione agarose beads at the end of the experiment. Green boxes indicate pulled-down effector proteins.

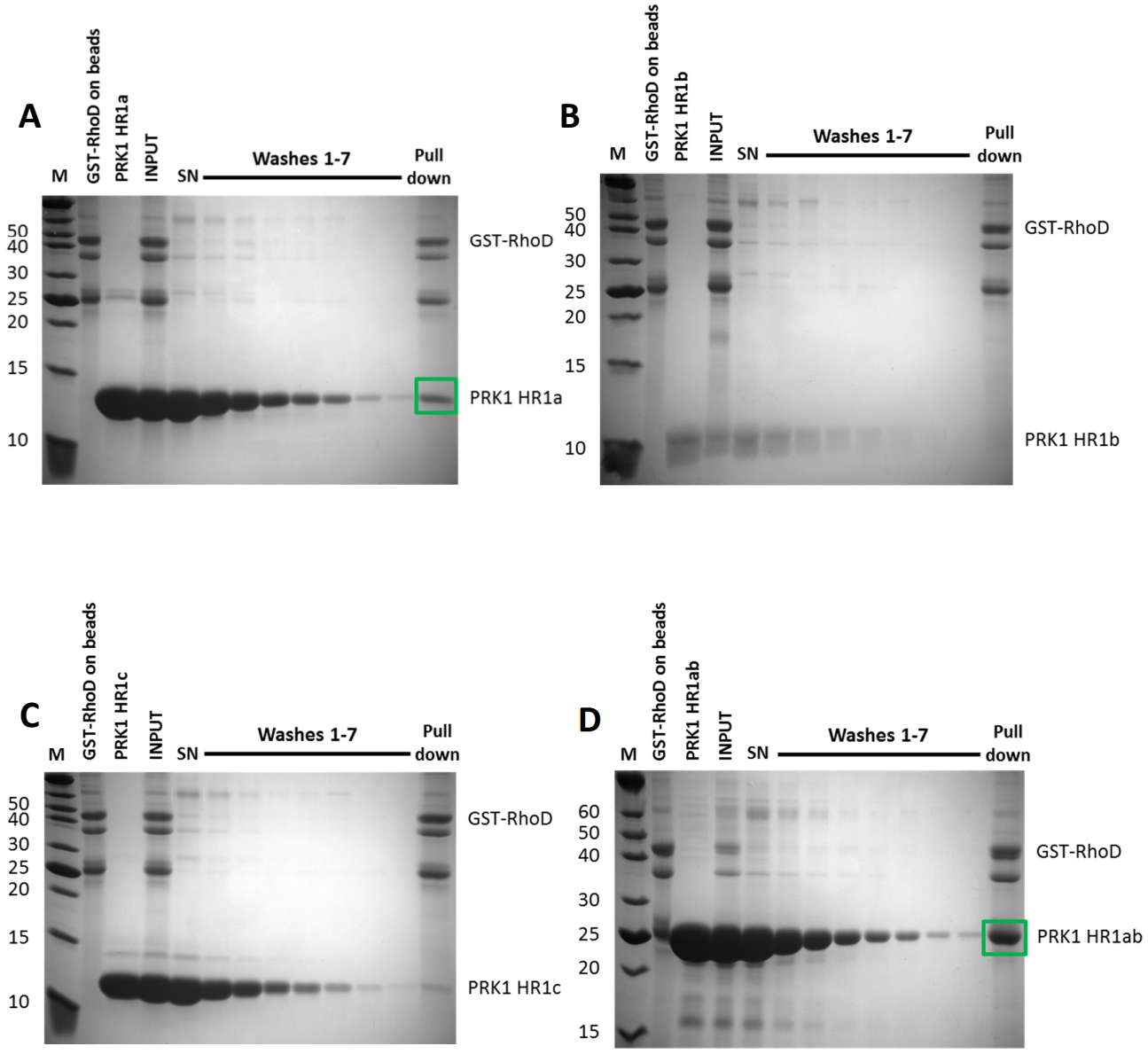


Figure 6.3 RhoD interacts with PRK1. 2-way pull-downs between GST-RhoD Δ4 F25N Q75L and (**A**) PRK1 HR1abc, (**B**) PRK1 HR1b, (**C**) PRK1 HR1c or (**D**) PRK1 HR1ab. HR1 effector protein was added to GST-RhoD immobilised on glutathione agarose beads. Following incubation, unbound protein was washed off. Binding was assessed by SDS-PAGE on gels of the appropriate percentage. M for molecular weight markers and SN for supernatant after incubation. The pull-down sample represents the proteins bound to the glutathione agarose beads at the end of the experiment. Green boxes indicate pulled-down effector proteins.

6.2.4 Expression and purification of GTPases

Untagged Rho GTPases were needed for use in experiments which would measure their affinity for HR1 domains as well as other biophysical studies (see section 6.3). Proteins were purified using glutathione agarose purification, the N-terminal GST tag was cleaved with HRV 3C protease (RhoA Δ4 F25N Q63L) or thrombin (Cdc42 Δ7Q61L and RhoA Δ12 F25N Q63L (see Chapter 5)) and the proteins were further purified using size-exclusion chromatography on a Superdex75 16/60 column. Figure 6.4 shows the purification of RhoA Δ4 F25N Q63L as an example.

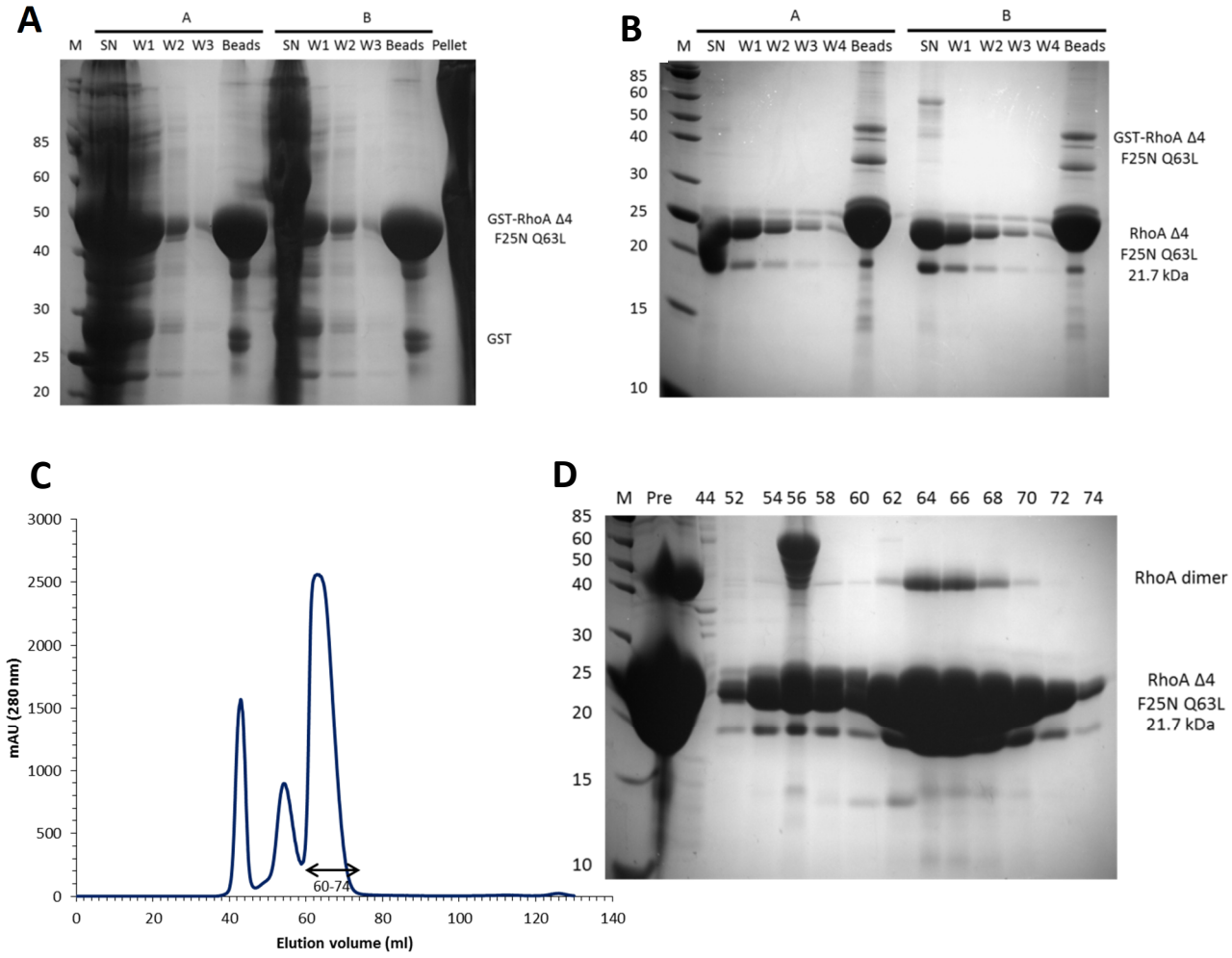


Figure 6.4 Purification of RhoA Δ4 F25N Q63L and corresponding SDS-PAGE analysis. (**A**) GST-RhoA Δ4 F25N Q63L immobilised on glutathione agarose beads. Gel was 12%. SN for supernatant, W1-4 for the 4 washes and Beads for 50% bead slurry. (**B**) RhoA Δ4 F25N Q63L washes post-cleavage. Gel was 15%. SN for supernatant and W1-4 for the 4 washes post-cleavage. (**C**) Gel filtration chromatogram of RhoA Δ4 F25N Q63L on an S75 column. Blue line indicates A_{280}. (**D**) SDS-PAGE analysis of chromatogram in (**C**) on a 15% Laemmli gel. 'Pre' for sample loaded on the column. The double arrows indicate the fractions pooled.

6.2.5 Nucleotide exchange

To lock the GTPase in its active, GTP-bound conformation, to allow binding to effectors, a slowly hydrolysable analogue, GMPPNP, was mixed with each GTPase in an exchange reaction. The protein was subsequently re-purified on a Superdex75 16/60 column to remove excess GMPPNP and purify the protein from any aggregates formed during the exchange. To check for successful nucleotide exchange, a small amount of protein was precipitated to release its nucleotide, the identity of which was confirmed by HPLC by comparing with guanosine nucleotide standards (Figure 6.5).

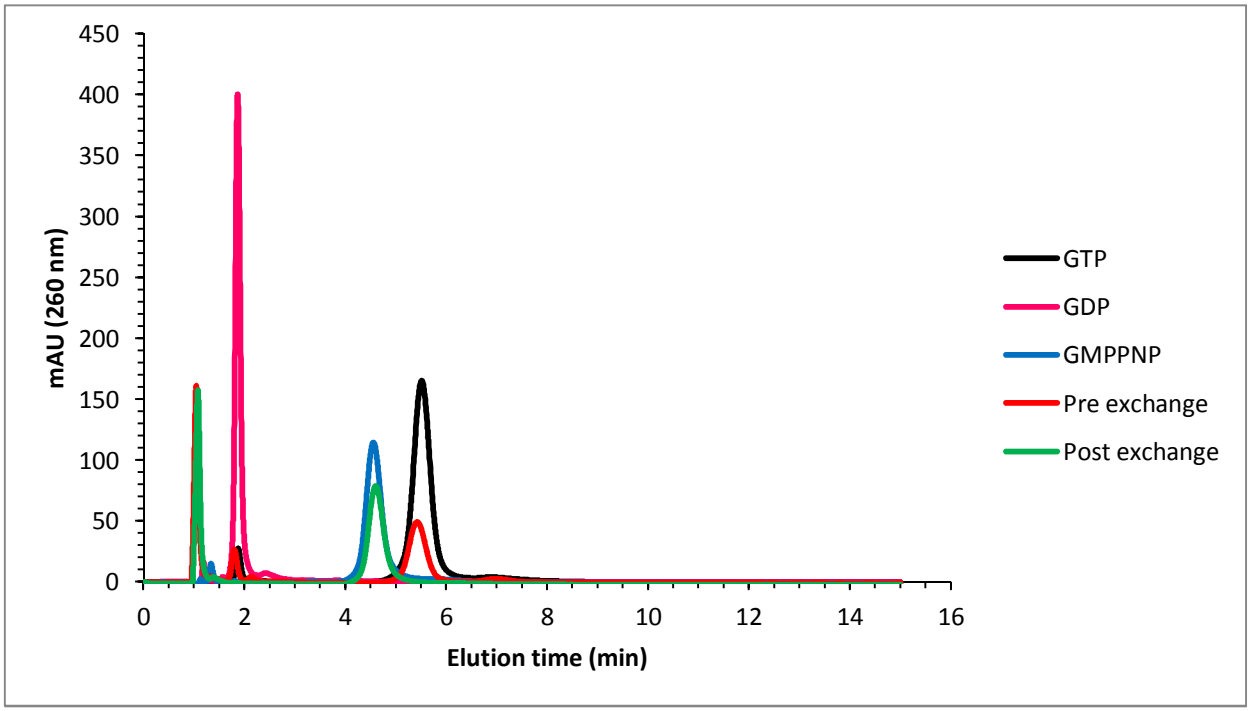


Figure 6.5 GMPPNP nucleotide exchange of GTPases. HPLC chromatogram of guanosine nucleotides. GDP, GTP and GMPPNP standards were compared to the nucleotide bound to RhoA Δ4 F25N Q63L. GTP in black, GDP in magenta and GMPPNP in blue. Nucleotide bound to RhoA Δ4 F25N Q63L before the exchange in red and after the exchange in green.

6.2.6 Measuring the affinity of Cdc42 for PRK3 HR1c

PRK3 HR1c had previously been shown to interact with Cdc42 (Sophocleous, 2014). To measure the affinity of this interaction by competition scintillation proximity assay (SPA), Cdc42 Δ7 Q61L was purified and labelled with [^{3}H]GTP. Cdc42/GST-ACK complexes (ACK for Activated-Cdc42 Kinase; a Cdc42 effector) were formed and PRK3 HR1c was used to compete Cdc42 off GST-ACK, resulting in a decrease in counts. Fitting the data gave a measure of the K_d at 8.08 μM (Figure 6.6). This affinity is about 10-100 times weaker than that of HR1a and HR1b for their cognate GTPases, suggesting that not all HR1 domains interact with Rho GTPases with similar affinities. Additionally, the affinity of Cdc42 for PRK3 HR1c is similar to the affinity of Cdc42 for the HR1 domains of TOCA/CIP4/FBP17. It is also important to note that PRK3 HR1c must be interacting with Cdc42 in a GTP-dependent manner as it competes with ACK. The number of counts remaining at the highest concentration of PRK3 HR1c was the same as when ACK competes with itself (Figure 6.6C). This indicates that HR1c must share the binding site on Cdc42 with ACK, i.e. the nucleotide-sensitive switch regions.

The binding of PRK1 HR1c to Cdc42 was also investigated by competition SPAs but it was found that it doesn't interact (not shown). This highlights differential binding of GTPases to the PRKs which may explain their non-redundant cellular roles.

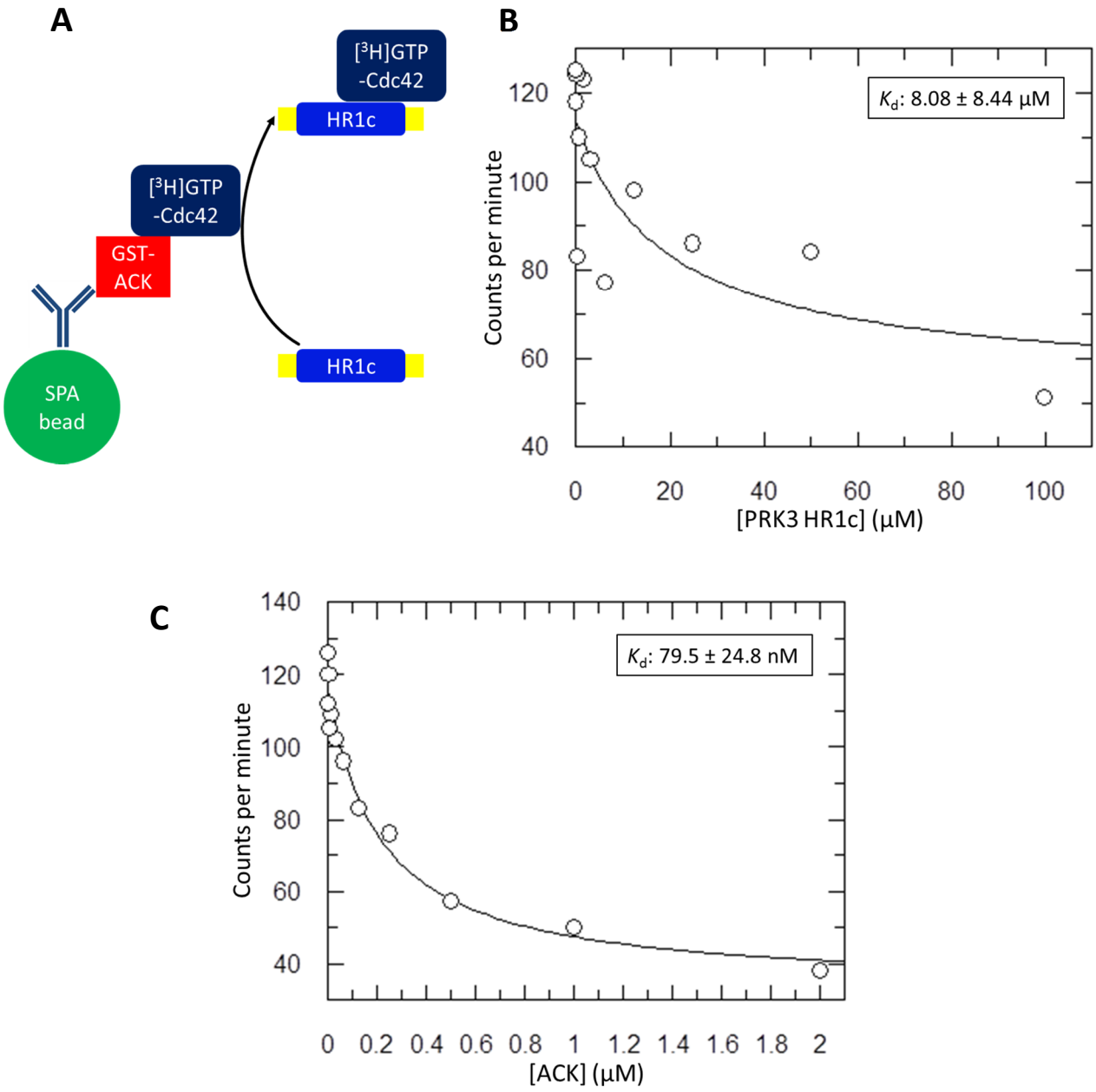


Figure 6.6 Competition SPA to measure the affinity of Cdc42 for PRK3 HR1c. (**A**) Principle of competition SPA. (**B**) Cdc42 Δ7 Q61L·[^{3}H·GTP] was mixed with GST-ACK and immobilised on SPA fluoromicrospheres via an anti-GST antibody. The GTPase was competed off with increasing concentrations of PRK3 HR1c, the data were fitted to a competition binding isotherm using non-linear regression and the K_d of the Cdc42/PRK3 HR1c interaction was calculated. (**C**) ACK was used to compete Cdc42 off GST-ACK as a positive control. ACK for Activated-Cdc42 Kinase.

6.2.7 Sequence alignments

To get an understanding of how RhoD may be interacting with PRK1 HR1a, the amino acid sequences of RhoA and RhoD were aligned (Figure 6.7). Out of the 12 interacting residues of RhoA with HR1a, 10 are fully conserved in RhoD. The remaining two residues are also weakly conserved. It is therefore very likely that HR1a binds to RhoD using contact II. Furthermore, none of the RhoA/HR1a contact residues are conserved in RhoD, which further implies that RhoD uses contact II exclusively to bind to PRK1 HR1a.

```
RhoA      ------------MAAIRKKLVIVGDGACGKTCLLIVFSKDQFPEVYVPTVFENYVADIEV 48
RhoD      MTAAQAAGEEAPPGVRSVKVVLVGDGGCGKTSLLMVFADGAFPESYTPTVFERYMVNLQV 60
                      ..   *:*:****.****.**:**:.. *** *.*****.*:.:::*

RhoA      DGKQVELALWDTAGQEDYDRLRPLSYPDTDVILMCFSIDSPDSLENIPEKWTPEVKHFCP 108
RhoD      KGKPVHLHIWDTAGQDDYDRLRPLFYPDASVLLLCFDVTSPNSFDNIFNRWYPEVNHFCK 120
          .** *.* :******:******** ***:.*:*:**.: **:*::** ::* ***:***

RhoA      NVPIILVGNKKDLRNDEHTRRELAKMKQEPVKPEEGRDMANRIGAFGYMECSAKTKDGVR 168
RhoD      KVPIIVVGCKTDLCKDKSLVNKLRRNGLEPVTYHRGQEMARSVGAVAYLECSARLHDNVH 180
          :****:** *.** :*:   .:* :   ***. ..*::**. :**..*:****: :*.*:

RhoA      EVFEMATRAALQARRGK-----KKSGCLVL  193
RhoD      AVFQEAAEVALSSRGRNFWRRITQGFCVVT  210
           **: *:..**.:*  :     .:. *:*
```

Figure 6.7 Sequence alignment of RhoA and RhoD. Sequences were aligned with Clustal Omega. Asterisks highlight identical residues at all positions, colons highlight conservation of a strong group and periods highlight conservation of a weak group. Residues which are involved in the RhoA/PRK1 HR1a interaction and are identical between RhoA and RhoD are highlighted in yellow, while non-identical residues are highlighted in cyan. RhoA/HR1a contact I residues are highlighted in pink.

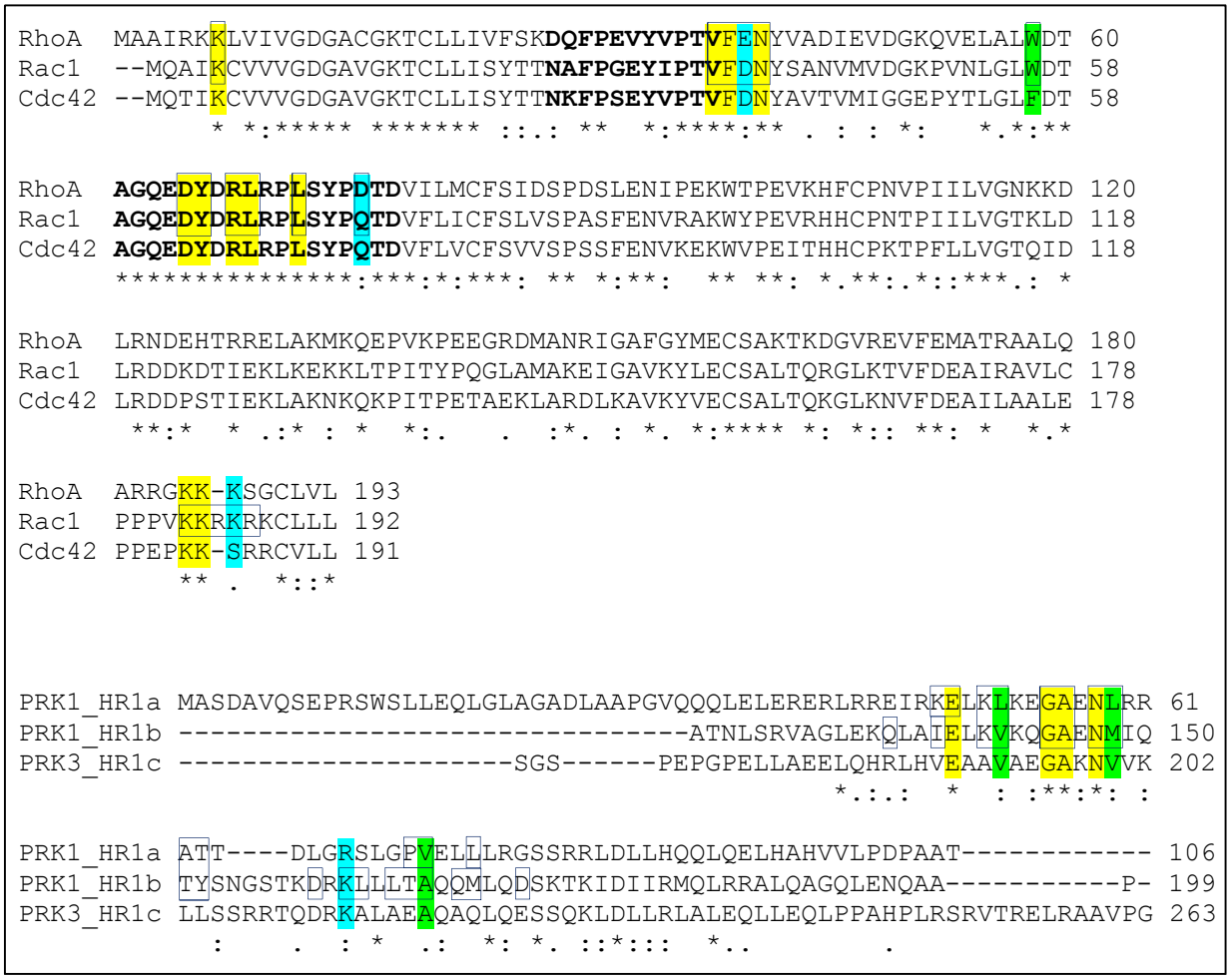


Figure 6.8 Sequence alignments of Rho family GTPases and their HR1 effector domains. (**A**)-(**B**) Sequence alignments were performed using Clustal Omega. Asterisks highlight identical residues at all positions, colons highlight conservation of a strong group and periods highlight conservation of a weak group. Interacting residues in the RhoA-PRK1 HR1a (Maesaki *et al.*, 1999) and Rac1-PRK1 HR1b (Modha *et al.*, 2008) structures are boxed. Interacting residues which are conserved are highlighted. Yellow for fully conserved residues, green for conserved non-polar or hydrophobic residues and cyan for conserved charged or polar residues. (**A**) Sequence alignment of human RhoA, Rac1 and Cdc42. The switch regions are in bold. (**B**) Sequence alignment of PRK1 HR1a, PRK1 HR1b and PRK3 HR1c.

Figure 6.8 highlights the residues in RhoA and Rac1 that are known to interact with PRK1 HR1a and PRK1 HR1b, respectively, and are also conserved in Cdc42 and PRK3 HR1c, suggesting that this is an interaction involving the switches, which is consistent with PRK3 HR1c competing ACK off Cdc42 in the SPA above. Most interacting residues are fully or strongly conserved between the HR1 domains and between the GTPases, which may suggest that additional features, such as HR1 rigidity or oligomerisation may influence GTPase specificity for HR1 domains.

6.3 How does oligomerisation affect GTPase binding?

6.3.1 RhoA binds to monomeric HR1a

Previous work in the lab showed that RhoA interacts with the PRK1 HR1a domain (Owen *et al.*, 2003; Hutchinson et *al.*, 2011). Scintillation proximity assays used to prove this interaction were performed with nM concentrations of HR1a. At such low concentrations, which are >1000 times below its dimerisation K_d, HR1a was monomeric. Therefore we know that monomeric HR1a can interact with RhoA, forming a 1:1 complex, and that HR1a uses the contact II binding interface of RhoA (Hutchinson *et al.*, 2011).

6.3.2 RhoA binds to dimeric HR1a

To test whether RhoA can bind to dimeric HR1a, AUC sedimentation velocity experiments were conducted (Figure 6.9). RhoA and HR1a were mixed at equimolar concentrations. The concentration chosen was 150 µM, which is the approximate dimerisation K_d of HR1a. At this concentration, the individual proportions of monomer and dimer are equal. This suggests that if RhoA can bind to both monomeric and dimeric HR1a, the complexes expected to be detectable are RhoA + HR1a (1:1 complex), RhoA + 2 x HR1a (1:2 complex) as well as unbound RhoA, unbound monomeric HR1a and unbound dimeric HR1a. The abundance of each of these five possible species depends on which interactions actually occur in solution, as well as their affinity. Figure 6.9A shows the sedimentation profile of RhoA which is the control experiment and provides the s-value of unbound RhoA, to help identify species present in solution in the experiment when HR1a is added. Figure 6.9B shows the sedimentation profile of RhoA mixed with HR1a at equimolar concentrations. From the five possible species, only two were detected. The first species was identified to be unbound RhoA, as the s-value is identical to that of free RhoA (Figure 6.9C). The interaction between the two proteins is confirmed because the abundance of the unbound RhoA species goes down despite the total RhoA concentration being equal between the two

experiments. The second species is therefore the bound species and the fit suggests a molecular weight of 49.6 kDa. As discussed in Chapter 3, the predicted MW is dependent on the fitted frictional ratio, which in turn is inversely proportional to the diffusion of the sedimenting species. In this case, there are two sedimenting species, so the frictional ratio is expected to be a global average representing both species. RhoA has a frictional ratio of 1.11 and the complex, which is the major species in this experiment, has a frictional ratio of 1.31, which suggests that this species is less globular than RhoA but slightly more globular than the dimerising HR1a (whose frictional ratio is 1.33). It is likely that this is a 1:2 complex of RhoA bound to dimeric HR1a. The reason why RhoA must be bound to dimeric HR1a rather than two HR1a monomers is that from the mutational analysis done in our lab, only one of the two contact sites is used by HR1a when bound to RhoA (Hutchinson *et al.*, 2011). This species is unlikely to be the 1:1 complex. If the fitted frictional ratio were to be lower, then the predicted MW would also be lower but for the MW to be 33 kDa (that of the 1:1 complex), the frictional ratio would have to be less than 1, i.e. the species would have to be more spherical than a sphere, which is not possible. The identification of the complex as the 1:2 complex is also supported by the presence of free RhoA which was detected, despite the equimolar concentrations used. To sum up, RhoA was found to stabilise the HR1a dimer and despite the published RhoA + HR1a structure, that species was not detected in solution at all.

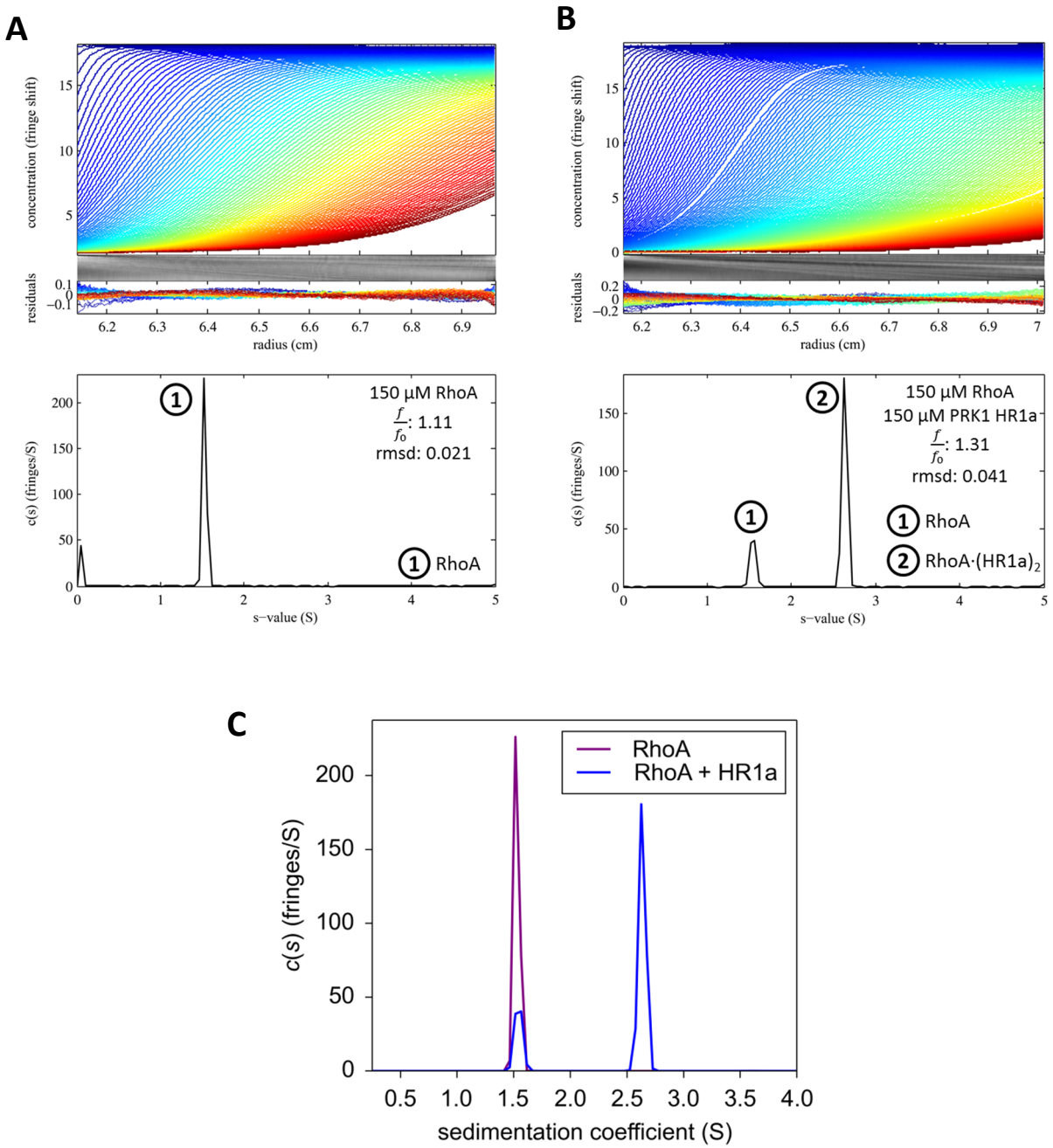


Figure 6.9 RhoA Δ4 F25N Q63L sedimentation velocity experiments. (**A**)-(**B**) Top panel represents the measured interference fringe distribution (blue to red from start to finish) and the corresponding fringe pattern. The middle panel represents the residuals of the fit, indicating how closely the experimental data match the fitted model. The grey image is a picture of the AUC cell which indicates the level of non-ideality in solution. The bottom plot represents the best-fit c(s) sedimentation coefficient distribution (against s-value (S)). Peaks represent protein species. % abundance was calculated by integrating the area under each peak. The fit rmsd is indicated. Sedimentation profiles of (**A**) 150 µM RhoA Δ4 F25N Q63L (**B**) RhoA Δ4 F25N Q63L and 150 µM PRK1 HR1a (**C**) Overlay of the c(s) distributions.

6.3.3 Measuring the affinity of dimeric PRK1 HR1a for RhoA

RhoA may bind to the HR1a dimer directly or via a step-wise mechanism, i.e. RhoA firstly binds to a single HR1a domain and subsequently to a second HR1a monomer, as summarised in the equation below.

$$\mathrm{RhoA} \rightleftharpoons \mathrm{RhoA{\cdot}HR1a} \rightleftharpoons \mathrm{RhoA{\cdot}(HR1a)_2}$$

To measure the affinity of RhoA for the HR1a dimer an isothermal titration calorimetry experiment (ITC) was conducted. RhoA at 1 mM concentration was placed in the syringe and injected into the cell which contained 100 µM PRK1 HR1a (60% monomer and 40% dimer according to AUC results in chapter 3). Heat changes were observed and the data were fit to a binding isotherm to calculate the K_d of the interaction (Figure 6.10). No step-wise binding was observed.

The K_d was calculated to be 602 nM, while the stoichiometry of the interaction as shown by N, was 0.643, suggesting that not all of the titrant is used to saturate binding to the protein in the cell. This is consistent with what was observed in AUC, i.e. RhoA forms a 1:2 complex with HR1a. The reaction was shown to be exothermic, with an unfavourable, negative change in entropy.

The affinity of RhoA Δ4 shown here is weaker than either RhoA Δ12 or RhoA full-length (Hutchinson *et al.*, 2013) but this is likely due to the different methods which were used.

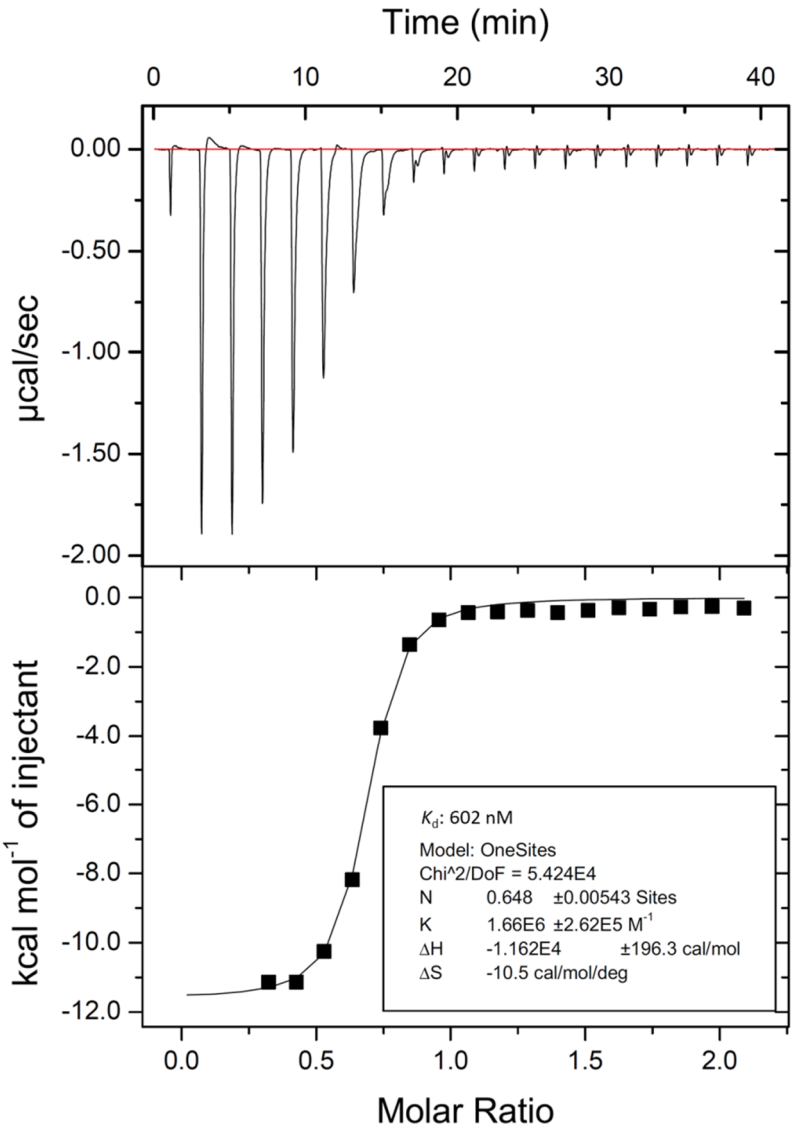


Figure 6.10 Measuring the affinity of RhoA Δ4 F25N Q63L for the PRK1 HR1a dimer. 1 mM RhoA Δ4 F25N Q63L·GMPPNP in the syringe was injected into 100 µM of PRK1 HR1a in the cell of an ITC instrument. Heat changes were recorded and the data were fit to a binding isotherm used to calculate the K_d, ΔH, ΔS and stoichiometry of the interaction.

6.4 Investigating the binding of RhoA to the HR1abc tri-domain

Previous studies investigating the interactions of HR1 domains with GTPases used isolated HR1 domains or the HR1ab di-domain. Currently it is not known whether the HR1abc tri-domain actually interacts with GTPases, and while very likely, this has never been proven. As shown in the previous Chapter, the three PRK1 HR1 domains are not independent of each other. As a result, it is possible that these intra- and inter-molecular interactions inhibit binding of GTPases or reduce the affinity of such interactions. Conversely, GTPases may preferentially bind to the HR1abc trimer similar to how RhoA binds and stabilises the PRK1 HR1a dimer (section 6.3.2). Additionally, cooperative binding of GTPases by two or more HR1 domains of HR1abc may also occur and has been observed in the case of the PRK1 HR1ab di-domain binding to RhoB more tightly than HR1a does (Hutchinson *et al.*, 2013).

6.4.1 2-way pull-downs between G proteins and HR1abc

To investigate whether HR1abc can interact with GTPases, 2-way pull-downs were carried out (Figure 6.11). Specifically, the binding of PRK1 HR1abc and PRK3 HR1abc to RhoA was tested. Both HR1abc tri-domains interacted with RhoA, indicating that the presence of all three HR1 domains does not interfere with binding. As *in vivo* studies have shown that binding of RhoA to the PRKs is necessary for their activation, it is perhaps unsurprising that the HR1abc tri-domain, which is the PRK regulatory region and physiologically more relevant, does interact with RhoA. These pull-downs do not indicate however if RhoA preferentially binds to a particular oligomeric state of HR1abc.

The HR1abc tri-domains were added to a final concentration of 100 μM. Most of the effector protein remained bound to GST-RhoA and did not dissociate during the washes. This suggests that the K_d of the interaction between RhoA and either HR1abc is well below 100 μM and that monomer, dimer and trimer all bind to RhoA (not necessarily with equal affinity).

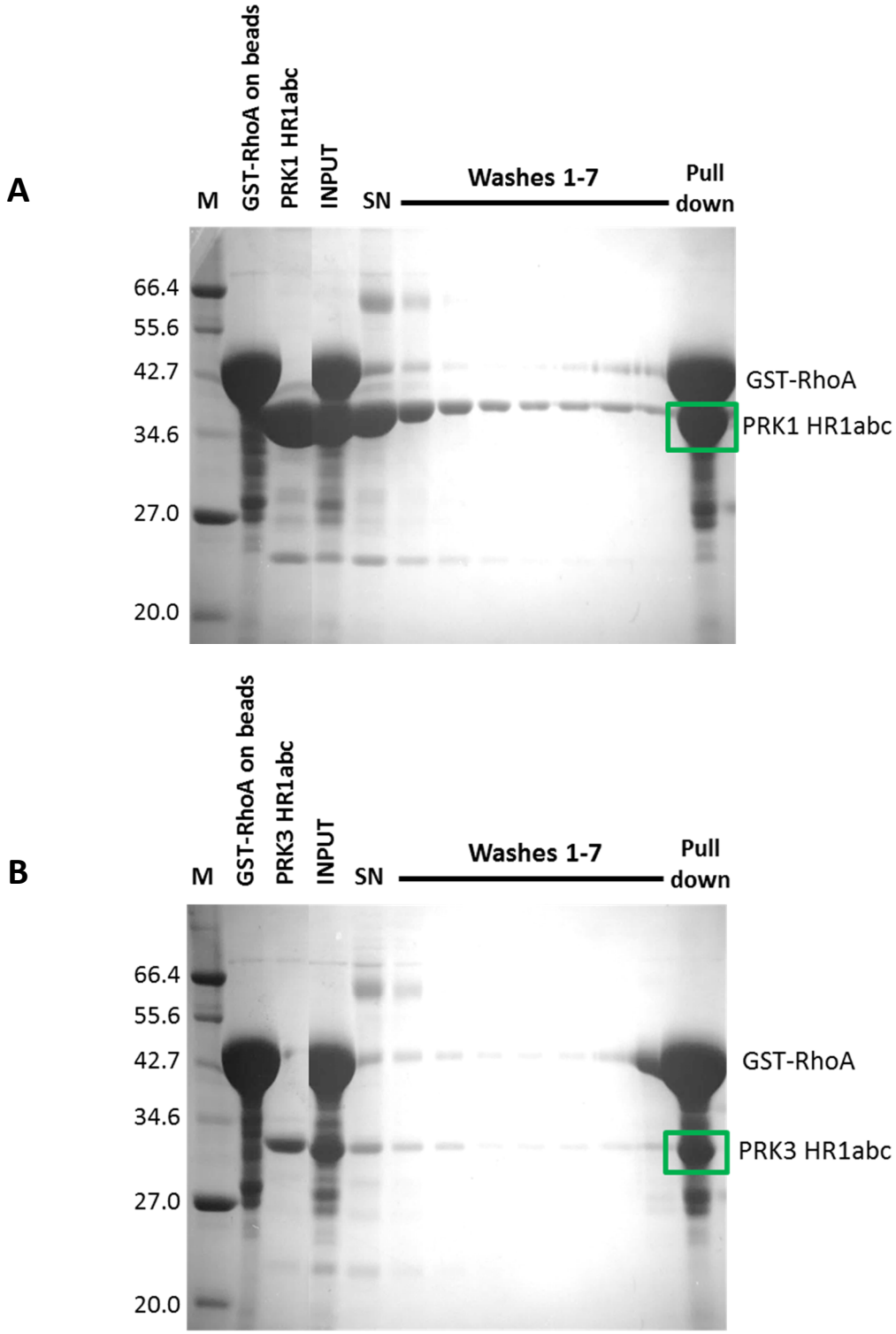


Figure 6.11 RhoA interacts with PRK1 HR1abc and PRK3 HR1abc. 2-way pull-downs between GST-RhoA Δ12 F25N Q63L and (**A**) PRK1 HR1abc or (**B**) PRK3 HR1abc. HR1abc effector protein was added to GST-RhoA immobilised on glutathione agarose beads. Following incubation, unbound protein was washed off. Binding was assessed by SDS-PAGE on 12% gels. M for molecular weight markers and SN for supernatant after incubation. The pull-down sample represents the proteins bound to the glutathione agarose beads at the end of the experiment. Green boxes indicate pulled-down effector proteins. (**C**) The nucleotide state of GST-RhoA was assessed by HPLC and compared to nucleotide standards.

6.4.2 Measuring the affinity of PRK1 HR1abc for RhoA

The affinity of PRK1 HR1abc for RhoA was determined by competition SPAs (Figure 6.12). RhoA/PRK1 HR1a complexes were formed and HR1a was competed off with HR1abc and the decrease in counts was used to calculate a K_d for the RhoA/HR1abc interaction. The K_d was determined to be 21.9 μM which is more than 100 times lower than the affinity of HR1a or HR1ab for RhoA.

The maximum concentration of HR1abc used in the assay was 60 μM. At this concentration HR1abc is expected to be mainly monomeric based on the AUC results on this protein (some dimer and trimer are still expected in solution). The binding seen here therefore suggests that HR1abc monomer does not interact tightly with RhoA. The monomer of the tri-domain interacts more weakly with RhoA than HR1a or HR1ab do, suggesting that intra-molecular, inter-domain interactions may be abrogating RhoA binding or that there is differential binding of dimers and trimers, which is not modelled in the equation used to fit the data.

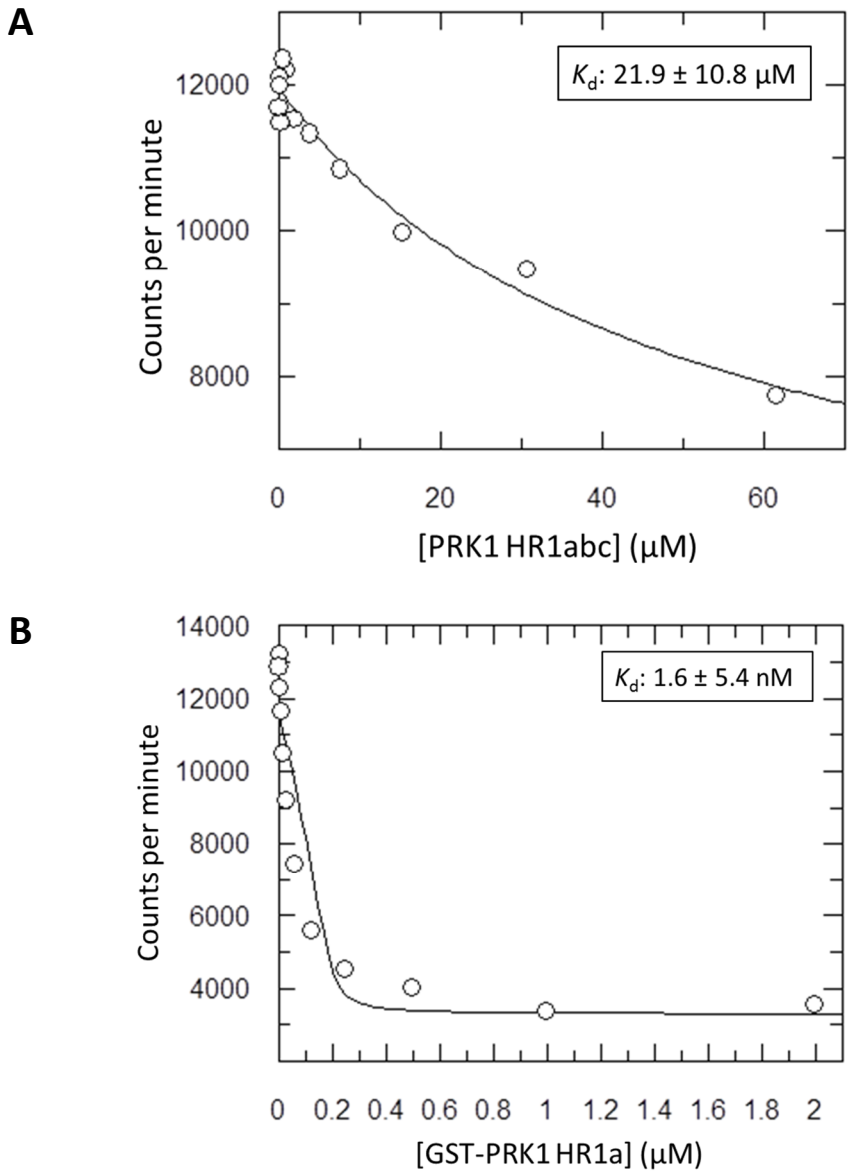


Figure 6.12 Competition SPA to measure the affinity of RhoA for PRK1 HR1abc. (**A**) RhoA Δ7 F25N Q63L·[^{3}H·GTP] was mixed with PRK1 HR1a-His_6 and immobilised on SPA fluoromicrospheres via an anti-His antibody. The GTPase was competed off with increasing concentrations of PRK1 HR1abc, the data were fitted to a competition binding isotherm using non-linear regression and the K_d of the RhoA/PRK1 HR1abc interaction was calculated. (**B**) GST-PRK1 HR1a was used to compete RhoA off PRK1 HR1a-His_6 as a positive control.

6.5 Conclusions

The PRKs form an effector alphabet for Rho GTPases. Figure 6.13 summarises all known interactions of PRK1 and PRK3 with Rho GTPases.

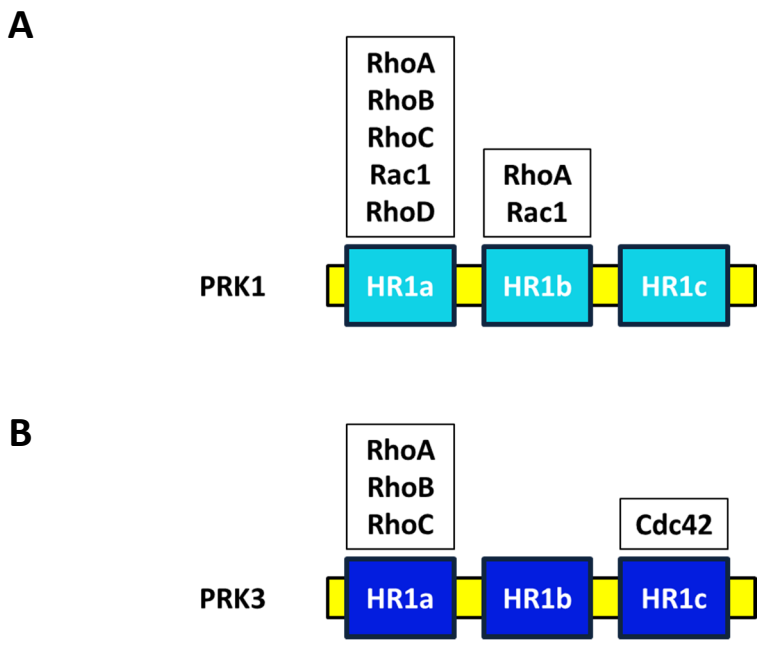


Figure 6.13 Summary of known PRK1 and PRK3 HR1 domain interactions with Rho GTPases. (**A**) PRK1 interactions (**B**) PRK3 interactions

In this chapter it has been shown using a combination of AUC and ITC that RhoA stabilises the PRK1 HR1a dimer to form a 1:2 complex. This result highlights the utility of both of these methods and challenges the naïve assumptions often made in binding studies where the stoichiometry of the interacting species is thought to be 1:1. This is not the first time RhoA has been shown to bind to a dimeric effector. Specifically, the ROCK effectors of RhoA bind to the GTPase switches via a parallel coiled coil which is formed when two long α-helices dimerise (Dvorsky *et al.*, 2004). As the HR1a dimer is stabilised in complex with RhoA, it is possible that all four helices make contact with RhoA. A step-wise model may also be possible, where initially only one coiled coil interacts with RhoA and subsequently its dimerisation K_d is lowered, allowing recruitment of a second HR1a to form the 1:2 complex.

Other examples of Ras superfamily GTPases binding to dimeric effectors include Rac1 (Rho family) binding to a dimer of Arfaptin (Tarricone *et al.*, 2001) and Arf6 binding to a dimer of

MKLP1 (Makyio *et al.*, 2012). Stabilisation of dimers of effectors by their cognate GTPase has been previously studied. Cdc42, another GTPase in the Rho family, has been shown to induce dimerisation of its effector IQGAP (LeCour Jr. *et al.*, 2016). This mechanism of dimer stabilisation is therefore shared across the family.

In Chapter 3 it was shown that PRK3 HR1a has a dimerisation K_d of 1 mM, which is approximately 6 times higher than the dimerisation K_d of PRK1 HR1a. PRK3 HR1a also interacts 100 times more weakly with RhoA than PRK1 HR1a does (Hutchinson *et al.*, 2013). It is possible that RhoA does not stabilise the PRK3 HR1a dimer or conversely that low levels of PRK3 HR1a dimer are not sufficient to form the 1:2 complex. This suggests, that GTPases may have distinct effects on the structure of full-length PRKs depending on their specific interactions with HR1 domains.

RhoA·GTP is required for activation of the PRKs. Mechanistically this has never been investigated but it has been speculated that RhoA activates the PRKs by binding to HR1a and thereby removing the autoinhibitory effect caused by the pseudosubstrate region in this domain (Kitagawa *et al.*, 1996). In this chapter, the characterisation of the RhoA/HR1a complex by biophysical methods such AUC and ITC has revealed the stabilisation of the HR1a dimer, which is otherwise in a reversible monomer-dimer equilibrium when free in solution. RhoA likely also induces formation of HR1abc dimers in the context of the full length kinase. This may have a critical role in the activation of the PRK1, e.g. by cross-phosphorylation as seen in the case of tyrosine kinase receptors or by enhancing the local concentration of catalytic domains which are activated by inter-molecular PIF motif interactions. These dimers are likely to be different to the inactive dimers seen in PRK2 (Bauer *et al.*, 2012).

Recent work in our lab by a Master's student under my guidance investigated the effect that RhoB bladder cancer mutations have on binding to PRK1. RhoB E29Q bound to HR1a and HR1ab with comparable affinity, suggesting that HR1b can no longer interact with RhoB. Glu29 is located in contact I of RhoB and mutating it to Gln reverts it back to the wild type RhoA, which cannot bind to the PRKs using contact I (Hutchinson *et al.*, 2011). This suggests that HR1ab may bind cooperatively to RhoB, with HR1a binding to contact II and HR1b binding to contact I. The reduced affinity of the mutant RhoB for PRK1 may result in the

disruption of RhoB/PRK1 endosomal signalling or the inability of PRK1 to localise to endosomes and may explain the tumour suppressor properties of RhoB.

Rac1 interacts with PRK1 HR1a with a similar affinity to the RhoA/HR1a interaction. Binding of Rac1 to dimeric HR1a has never been investigated however. It therefore remains to be seen if Rac1 can also stabilise the HR1a dimer or if this effect is exclusive to RhoA. Assuming that RhoA activates PRK1 through HR1a-mediated dimerisation, then Rac1 may also bind to dimeric HR1a and activate PRK1 in a similar manner (Lu & Settleman, 1999).

PRK3 HR1c is no more like the TOCA family HR1 domains than other PRK HR1 domains, yet it interacts with Cdc42 with near-identical affinity to the TOCA family HR1 domains. The melting temperatures of the HR1 domains are also different, suggesting that it is not the rigidity of these domains which makes them all interact with Cdc42. Put together, these data could be indicative of specificity being defined by sequence and structural characteristics of the G protein rather than the HR1 domain. This is further highlighted by the oligomeric state of these effector proteins; the TOCA family HR1 domains are monomeric while PRK3 HR1c exists in a monomer-dimer-tetramer equilibrium. Whether Cdc42 can stabilise the PRK3 HR1c dimer (or tetramer) like RhoA stabilises the PRK1 HR1a domain is also possible but has not currently been investigated. If Cdc42 does bind to the oligomeric PRK3 HR1c, then this may explain why PRK1 HR1c, which is monomeric, does not interact with Cdc42.

It is also likely that TCL and TC10, which belong to the Cdc42 subfamily, can also interact with PRK3 HR1c. It may also be the case that PRK1 HR1c, which cannot interact with Cdc42, interacts with either of these two-related proteins. TC10 is involved in signalling of GLUT transporters to the membrane and PRK1 is also known to regulate this process. TC10 may therefore be acting via PRK1 to control this signalling.

Interestingly, the SH3 domain of Graf and Graf2 interacts exclusively with the proline-rich region of PRK3 (Shibata *et al.*, 2001). Graf and Graf2 contain a RhoGAP domain which can stimulate the intrinsic GTP hydrolysis of RhoA and Cdc42. It is possible therefore that PRK3 recruits these RhoGAPs to revert the GTPases back to their inactive, GDP-bound conformation to terminate signalling. The exclusive binding of Graf and Graf2 to PRK3 may

therefore be viewed as a feedback mechanism for the equally exclusive binding of PRK3 to Cdc42.

RhoD is considered to be one of the 'typical' Rho GTPases yet this view can be challenged as RhoD is set apart from other 'typical' GTPases due to its fast-cycling nature. It localises to early endosomes, where it regulates endosomal trafficking and dynamics via numerous effectors, including the hDia2C Diaphanous protein (Gasman *et al.*, 2003), c-Src kinase and the Rab5 effector, Rabankyrin. Specifically, RhoD recruits hDia2C from the plasma membrane (Gasman *et al.*, 2003), which suggests that a similar event may occur with RhoD and PRK1. A physiological role of PRK1 in early endosomes has not been reported yet, however RhoD may have a role in guiding PRK1 to late endosomes. At late endosomes, RhoB can then bind to PRK1 to regulate epidermal growth factor receptor trafficking (Gampel *et al.*, 1999). RhoD may therefore have a role in regulating how many PRK1 molecules localise to late endosomes, thereby indirectly regulating late endosomal signalling.

More Rho GTPases may interact with the PRKs. For example, eleven out of twelve amino acids involved in the RhoA/HR1a interaction are conserved in the three Rnd proteins (Figure 6.14), suggesting that the Rnd proteins may also interact with PRK1. Screening HR1 domains for binding against each Rho GTPase is imperative in identifying such interactions.

Figure 6.14 Sequence alignment of Rho GTPases. Sequences were aligned with Clustal Omega. Asterisks highlight identical residues at all positions, colons highlight conservation of a strong group and periods highlight conservation of a weak group. Residues which are involved in the RhoA/PRK1 HR1a interaction and Rac1/PRK1 HR1b interactions are highlighted in yellow. The triple-proline motif of Rac1 is highlighted in red.

Chapter 7 Discussion and future directions

The principal aim of this study was to obtain biochemical, biophysical and structural insight into the HR1 domains. This was been successful with PRK1 HR1 domains and to a lesser extent with the PRK3 HR1 domains. Chapter 3 provided extensive characterisation of the secondary structure of PRK1 and PRK3 HR1a, HR1b, HR1c, HR1ab and HR1abc proteins, as well as insight into the secondary structure and stability, for the first time, of HR1c domains. Furthermore, work in Chapter 3 revealed that PRK1 HR1a as well as all three PRK3 domains can all oligomerise, identifying HR1 domains for the first time as drivers of PRK oligomerisation. Chapter 4 presented the NMR solution structure of the PRK1 HR1c domain. This domain is structurally more diverse from HR1a and HR1b. While it conserves two helices which form a coiled coil, it also has a shorter, C-terminal helix which forms a triple coiled coil structure. Chapter 5 explored the inter-molecular interactions between PRK1 HR1 domains which were indicated by AUC experiments on the HR1abc tri-domain in Chapter 3. The key finding was that the three HR1 domains all interact with each other and it is these interactions which are thought to drive formation of a higher order complex (as observed for the HR1abc trimer). Chapter 6 then expanded this linked equilibrium model by directly connecting RhoA to PRK1 oligomerisation while also investigating novel GTPase interactions with HR1 domains.

The aim of this chapter will be to collectively review these findings and understand their significance in PRK signalling in the context of a simple, yet coherent model. The chapter also aims to highlight key questions which have arisen from this study but remain unanswered and provide future directions for this project.

7.1 Structural and biochemical insight into the HR1 domains

7.1.1 Intra- and inter-molecular interactions

HR1 domains are approximately 10 kDa each and despite their sequence identity being very low (~30%), they share structural conservation as they fold into coiled coils. The stability of HR1 domains however is one of their key differences as indicated by melting temperatures ranging from 33.2 to 60.1 ^{o}C (Hutchinson *et al.*, 2013; Chapter 3). Out of the three PRK1 domains, HR1a has the lowest melting temperature, and coincidentally is the only PRK1 HR1 domain which can oligomerise. Low melting temperature may therefore be important both in its dimerisation, but also in its ability to interact with HR1b and HR1c as well as with 5 Rho GTPases (RhoA, RhoB, RhoC, Rac1 and RhoD). In contrast, HR1b is thermodynamically more rigid and while it shares the contacts HR1a makes with RhoA in its interaction with Rac1, it only binds with nM affinity due to additional contacts made with the polybasic region in the C-terminus of Rac1 (Owen *et al.*, 2003; Modha *et al.*, 2008). HR1c is also more rigid than HR1a, and like HR1b it is monomeric in solution.

The PRK3 HR1 domains differ greatly from their PRK1 homologs. Firstly, both HR1a and HR1b have significantly lower melting temperatures than HR1c (HR1a: 33.2 ^{o}C, HR1b: 42.9 ^{o}C; 58 ^{o}C), and all three domains can oligomerise. Strangely, the domain with the highest melting temperature is the domain with the lowest dimerisation K_d, suggesting that rigidity cannot entirely account for HR1 oligomerisation. PRK3 HR1c has high sequence identity to PRK1 HR1c, and while PRK1 HR1c is monomeric, PRK3 HR1c forms a monomer-dimer-tetramer mixture.

The structure of PRK1 HR1c has been solved by solution NMR, revealing a coiled coil domain which has an additional C-terminal helix. This helix packs against the other two helices, forming a triple helix. This is an additional difference between HR1c and other HR1 domains, however the role of this C-terminal helix still remains unknown.

PRK1 HR1c still has not had a cognate GTPase identified. However it may have evolved to interact intra- and inter-molecularly with HR1a and HR1b, thereby promoting formation of HR1abc oligomers.

Figure 7.1 illustrates a linked equilibrium between HR1abc monomer, dimer and trimer and their RhoA-bound states. RhoA stabilises the HR1a dimer but whether it can stabilise the HR1abc dimer as well has not been investigated. If other GTPases are brought into this linked equilibrium e.g. if Rac1 can interact with HR1abc at the same time as RhoA does, then the system gets exponentially complicated. This leads to the question as to why is there a need to exist in so many different states, and not just two oligomeric states (one catalytically active and one catalytically inactive). The answer may have to do with the ability of each of these states to differentially interact with GTPases, leading to swift change in cellular localisation which is a property the PRKs are likely to have given their tremendously diverse cellular roles in multiple cellular compartments.

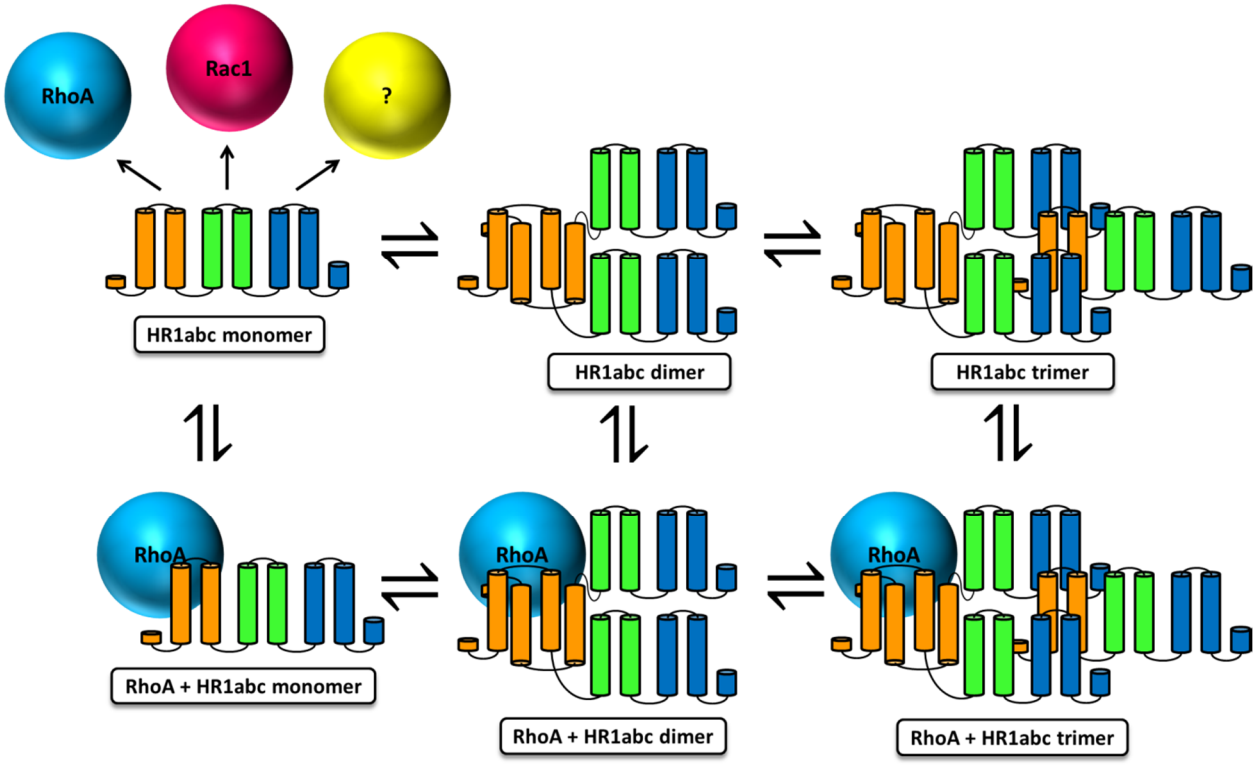


Figure 7.1 Model of RhoA binding to PRK1 HR1abc. Model of PRK1 oligomers and RhoA linked equilibria. Orange for HR1a, green for HR1b and blue for HR1c. GTPases are illustrated as spheres.

7.1.2 Wider implications for PRK signalling in cells

The findings in this study, in conjunction with published studies, allow us to delineate the PRK1 pathway in cells (Figure 7.2). PRK1 in the cytosol is inactive as it needs a series of activating factors which are only found in the plasma membrane. When the plasma membrane is enriched in particular phospholipids due to upstream signalling, PRK1 gets recruited via its C2-like domain which binds to these membrane phospholipids. Some activation of PRK1 may occur directly via arachidonic acid, PIP_2, PIP_3 and cardiolipin (Mukai *et al.*, 1994; Morrice *et al.*, 1994; Kitagawa *et al.*, 1995; Peng *et al.*, 1996; Yu *et al.*, 1997; Mukai & Ono, 2006; Palmer *et al.*, 1995), but PRK1 remains largely in an autoinhibited state (Kitagawa *et al.*, 1996; Yoshinaga *et al.*, 1999). RhoA also attaches to the plasma membrane via its geranylgeranylated C-terminus. Rho GEFs can then catalyse nucleotide exchange and as a result RhoA·GTP can interact with the HR1a domain of PRK1. RhoA therefore relieves PRK1 autoinhibition exerted by the HR1a pseudosubstrate region (Kitagawa, *et al.* 1996). RhoA also promotes dimerisation of PRK1 by stabilising the HR1a dimer (Chapter 6). Rac1, which can also activate PRK1 (Lu & Settleman, 1999), may also promote HR1a dimerisation, while simultaneous binding to HR1b may promote intra- and inter-molecular HR1 interactions, which can lead to HR1abc trimer formation. Dimerisation (or oligomerisation) may have additional activatory roles, e.g. leads to autophosphorylation of PRK1 which can activate its catalytic activity (Yoshinaga *et al.*, 1999). RhoA also enhances the activation of PRK1 by PDK1 (also localised to the plasma membrane via its lipid-binding pleckstrin homology domain) which phosphorylates the turn motif of the PRK1 kinase domain. A ternary complex between RhoA, PRK1 and PDK1 may form to cooperatively enhance these three events: removal of autoinhibition, phosphorylation and dimerisation. Following PRK1 activation, RhoA and PDK1 dissociate via a mechanism which is currently unknown.

PRK1 dimerisation by the HR1a domain brings all the HR1 domains of two PRK1 protomers into proximity. This increase in effective molarity promotes formation of intra- and inter-molecular HR1 interactions which involve the HR1c domain. Formation of the HR1abc trimer may be what causes dissociation of RhoA, as it binds more weakly to RhoA than HR1a does (Chapter 6). Activated, oligomerised PRK1 at the plasma membrane can now mediate

downstream signalling in numerous cellular processes, including regulation of the cytoskeleton, cell adhesion and regulation of pyrin inflammasome formation.

Activated PRK1 may then be localised to early endosomes by RhoD which binds to the HR1a domain. RhoB binds to PRK1 with greater affinity than other GTPases, suggesting that RhoB may then compete with RhoD for binding to PRK1, localising it to late endosomes (Mellor *et al.*, 1998). Here, PRK1 regulates the trafficking of epidermal growth factor receptors (EGFR), which may prolong EGFR signalling (Gampel *et al.*, 1999). When RhoB is active in the endosomes, it is thought to suppress Rac1 activity (Marcos-Ramiro *et al.*, 2016). Mechanistically this could involve PRK1, which can bind (at least separately) to both of these GTPases. When Rac1 is more active than RhoB, PRK1 may be translocated to the plasma membrane again, terminating RhoB/PRK1-dependent signalling in late endosomes.

PRK1 can also bind to the transactivation domain of the androgen receptor (AR), which may localise it to the nucleus. PRK1, in complex with the AR, bind to androgen responsive elements of AR-regulated genes. PRK1 phosphorylates Thr11 of histone H3 (Metzger *et al.*, 2008) while it can also phosphorylate three histone HDACs, events which are both linked to the regulation of chromatin remodelling.

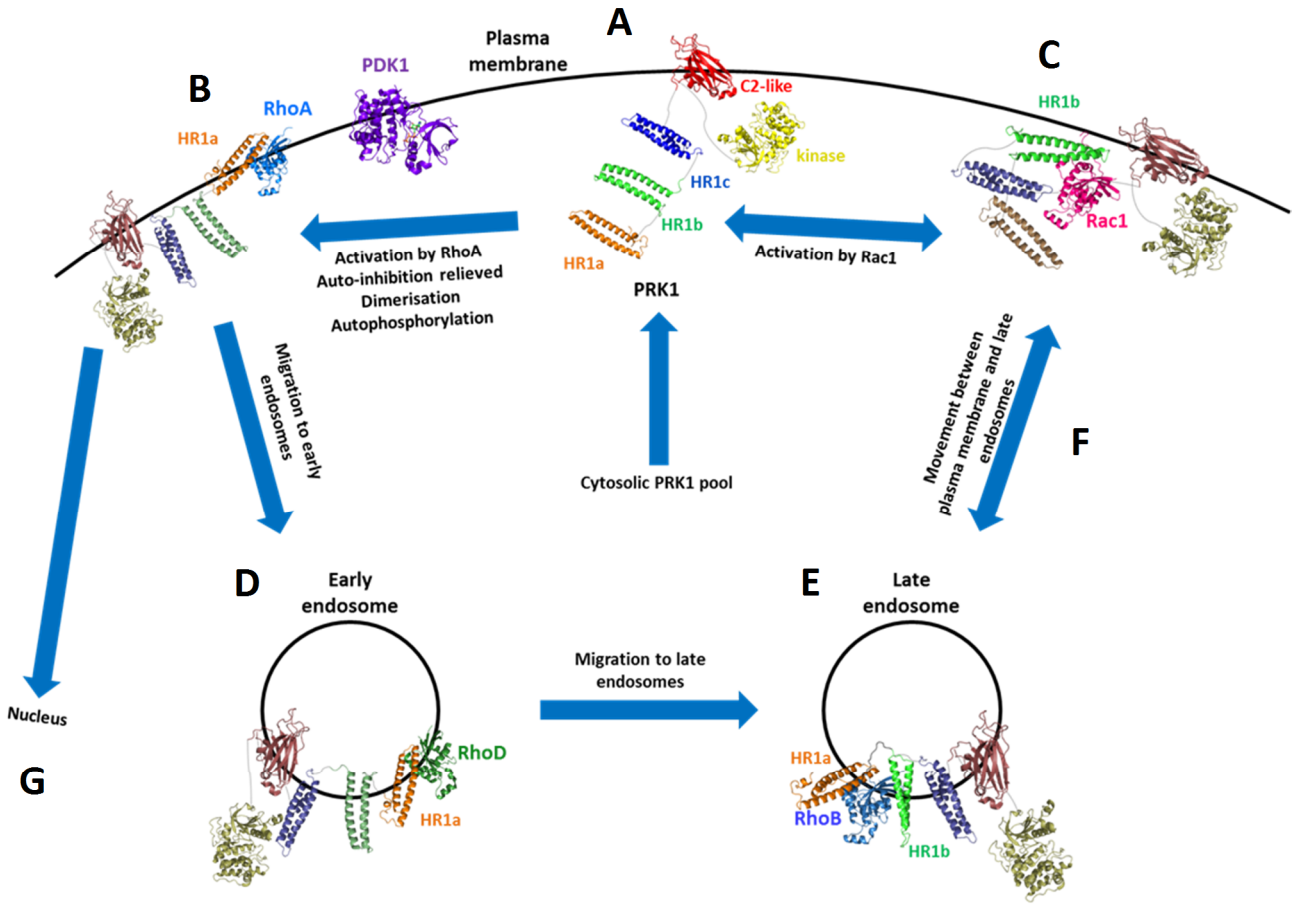


Figure 7.2 PRK1 activation and localisation pathways. Structures of PRK1 HR1a (orange), PRK1 HR1b (green), PRK1 HR1c (blue), PKC C2-like domain (red), PRK1 kinase domain (yellow), RhoA (marine blue), Rac1 (magenta) and PDK1 (purple) are shown in cartoon representation. The structures of RhoD (dark green) and RhoB (sky blue) in complex with HR1 domains are models and are also shown in cartoon representation. The proteins and HR1 domains relevant at each stage are highlighted in their respective colours, while all other domains are shadowed out. (**A**) Cytosolic PRK1 localises to the membrane via its C2-like domain in response to lipids. (**B**) PDK1 and RhoA both activate PRK1; the former via phosphorylation of the PRK1 kinase domain and the latter by relieving PRK1 autoinhibition and promoting dimerisation of the HR1a domain. Dimerisation enables further oligomerisation and may also enhance autophosphorylation, which may further activate PRK1 and release RhoA and PDK1. (**C**) PRK1 can also be activated by Rac1 which binds to the HR1b domain. Rac1 also binds to the HR1a domain which may also activate PRK1 via dimerisation. (**D**) PRK1 translocates to early endosomes, where it interacts with RhoD via its HR1a domain. (**E**) PRK1 then translocates to late endosomes, where it can interact with RhoB which may possibly involve HR1a and HR1b interacting with contact II and contact I, respectively. (**F**) PRK1 may relocate to the plasma membrane as a result of antagonistic RhoB/Rac1 signalling. (**G**) PRK1 also localises to the nucleus, possibly as a complex with the androgen receptor.

7.2 Future work

7.2.1 Characterisation of the dimerisation interface

A key question is about the dimerisation interface of PRK1. Previous attempts in our lab to study the HR1a domain by NMR have been unsuccessful due to extensive linewidth broadening, which can now be attributed to the chemical exchange between monomer and dimer. The ^{15}N HSQC PRK1 HR1ab spectrum at a concentration ten-fold lower than the dimerisation K_d of HR1ab resulted in the appearance of many new backbone amide peaks from the HR1a domain. However approximately 50-60 more peaks were still missing and it is probable that it is these peaks which are involved in dimerisation. Without a way of assigning the peaks which have appeared, it is impossible to assign the peaks which remain disappeared. Given more time, pH titrations could be carried out in an attempt to see if the dimerisation could be stabilised or stopped or if the rate of chemical exchange could be significantly lowered.

Hydrogen/deuterium exchange (H/D exchange) mass spectrometry (MS) is a method which has been used to map protein-protein, protein-lipid and other interactions (reviewed in Cao *et al.*, 2013). Proteins of interest are incubated with D_2O and are therefore labelled via exchangeable hydrogens. The reaction is then quenched, the protein is digested and peptides are analysed by MS. This could be used to identify the HR1a dimerisation site as this site would be less prone to H/D exchange.

A more laborious approach would involve the systematic alanine scanning of each HR1a residue and the characterisation of the oligomeric state of each mutant protein. Similar experiments may be used to investigate the dimerisation interfaces of the PRK3 HR1 domains and in particular PRK3 HR1c, which has a dimerisation affinity similar to that of PRK1 HR1a. The oligomerisation of the PRK2 HR1 domains has not been investigated due to the difficulty of obtaining soluble protein for HR1c and HR1abc proteins. Co-expressing with chaperones may solve this problem and then the oligomeric state of PRK2 domains could be investigated by AUC.

Yoshinaga *et al.* (1999) reported possible autophosphorylation sites on PRK1. One of these is Thr63 which is located in helix 1 of HR1a. Thr63 forms a Van der Waals contact with Trp58 of RhoA as seen in the RhoA/HR1a structure (Maesaki *et al.*, 1999). The distance between these two residues is measured to be 3.8 Å. This suggests that if Thr63 is phosphorylated, then the negatively charged phosphate would displace RhoA. This may be a regulatory mechanism, which may have a role in the dissociation of RhoA from PRK1 once PRK1 is activated. A phosphomimic mutant HR1a, e.g. Thr63Glu could be designed and the affinity for RhoA could be measured. A decreased affinity would suggest that upon PRK1 activation, autophosphorylation occurs which leads to RhoA dissociation.

7.2.2 Characterisation of the HR1abc tri-domain

The HR1abc trimer is likely to involve interactions between HR1 domains of different protomers which are not necessarily of 1:1 stoichiometry. Therefore had more time been available, AUC experiments of different combinations of HR1 domains would have been carried out. The information from these experiments would couple exceptionally well with the HR1 mapping data (Chapter 5), e.g. it could indicate if HR1a binds to more than one contact site on HR1c at a time, or if several molecules bind to the different sites.

Small Angle X-ray Scattering (SAXS) is a powerful method for obtaining structural information about proteins (Svergun, 2007). While SAXS strictly requires a monodisperse sample, size exclusion chromatography (SEC)-SAXS has recently allowed the structural characterisation of oligomers in a polydisperse sample (Fekry *et al.*, 2017). This is because the data obtained from each species as it elutes from a gel filtration column can be analysed separately. Therefore, (SEC)-SAXS could be used to study HR1abc oligomers.

The binding site of HR1a on HR1c has been determined in Chapter 5. However the reciprocal binding site on HR1a has not been determined. A way to investigate if the dimerisation site overlaps with the RhoA-interaction site, RhoA could be titrated into ^{15}N HR1c bound HR1a and competition HSQC titrations carried out to see if the peaks of interacting residues would return to their original chemical shift.

7.2.3 PRKs as Rho GTPase effectors

7.2.3.1 Identification of novel GTPase interacting partners

The results presented in this study demonstrate that many more GTPases may interact with the PRKs and in particular the HR1c domain but remain unknown. To identify new interacting GTPases, co-immunoprecipitations from cells could be carried out. This would require mutations to stop RhoA/Rac1 binding (as these two GTPases may bind more strongly than any other GTPases bind to PRK1). This could have a structural effect, e.g. loss of HR1-HR1 interactions, that may prevent other G proteins from binding to HR1c. Two-way pull-downs using purified proteins may therefore be a better option (although more laborious).

Additionally, the interaction of Cdc42 with PRK3 HR1c has only been seen *in vitro* and may therefore not be physiological. This interaction can therefore be confirmed by co-immunoprecipitations.

7.2.3.2 Differential use of contacts I and II

Based on alanine scanning of RhoB contact sites done by a student under my guidance, it is speculated that RhoB binds to both HR1a and HR1b using contact sites II and I, respectively. The binding site of HR1b for RhoB could be determined by carrying out NMR titrations.

It can be argued that RhoB may bring the HR1a and HR1b domains of two separate protomers into proximity, thereby mediating PRK oligomerisation or it may simply bind to HR1a and HR1b from a single protomer. It would be interesting to determine the oligomeric complexes formed between RhoB and HR1ab and if and how they differ from those of RhoA and HR1ab. Therefore AUC and SEC-SAXS experiments could be carried out. These results could highlight differences in the supercomplexes formed (as a result of differential use of the contact sites). This would explain both *in vivo* and *in vitro* evidence that RhoB is the cognate GTPase of PRK1 (as well as PRK2 and PRK3), as RhoB binds to PRK1 with higher affinity than RhoA and RhoC do (Hutchinson *et al.*, 2013) and it localises with PRK1 in endosomes to mediate EGFR signalling (Mellor *et al.*, 1998; Gampel *et al.*, 1999).

7.2.3.3 Simultaneous GTPase protein binding

Another key question which needs to be answered is whether two or more GTPases can simultaneously interact with a single PRK. For this, 3-way pull-downs can be carried out, using a GST-tagged GTPase, adding HR1abc and then adding another GTPase that is untagged.

Simultaneous GTPase binding would provide clues as to whether there is signalling convergence in cells, e.g. two GTPases (RhoA and Rac1 for example) can activate a PRK at the same time.

7.2.3.4 Characterisation of supercomplexes

RhoA has been shown to interact with the HR1a dimer, while the NMR solution structure of Rac1 with HR1b shows that Rac1 binds to monomeric HR1b. The effect of RhoB, RhoC, RhoD and Cdc42 on the oligomerisation of HR1 domains has not been investigated. Therefore AUC experiments could be performed to determine this.

As has been described previously, some Rho GTPases can interact with more than one HR1 domain. RhoA binds to PRK1 HR1a with higher affinity than it does to HR1b, while Rac1 binds to both HR1a and HR1b with similar affinities. The affinities of RhoA and Rac1 for HR1ab suggest that only contact II is used to bind to HR1 domains, so only one of these HR1 domains may bind at a given time. Assuming that the PRKs can interact simultaneously with more than one GTPase, different supercomplexes can potentially form in solution. These include two RhoA molecules each binding to HR1a and HR1b; two Rac1 molecules each binding to HR1a and HR1b; RhoA binding to HR1a and Rac1 to HR1b. In contrast, RhoB binds to HR1ab with higher affinity than both RhoA or Rac1 and it is speculated to bind to HR1b via contact I and to HR1a via contact II. This suggests that it could prevent binding of either RhoA or Rac1. Furthermore, bigger supercomplexes may form if multiple GTPases can interact with HR1 domains in an oligomerised PRK.

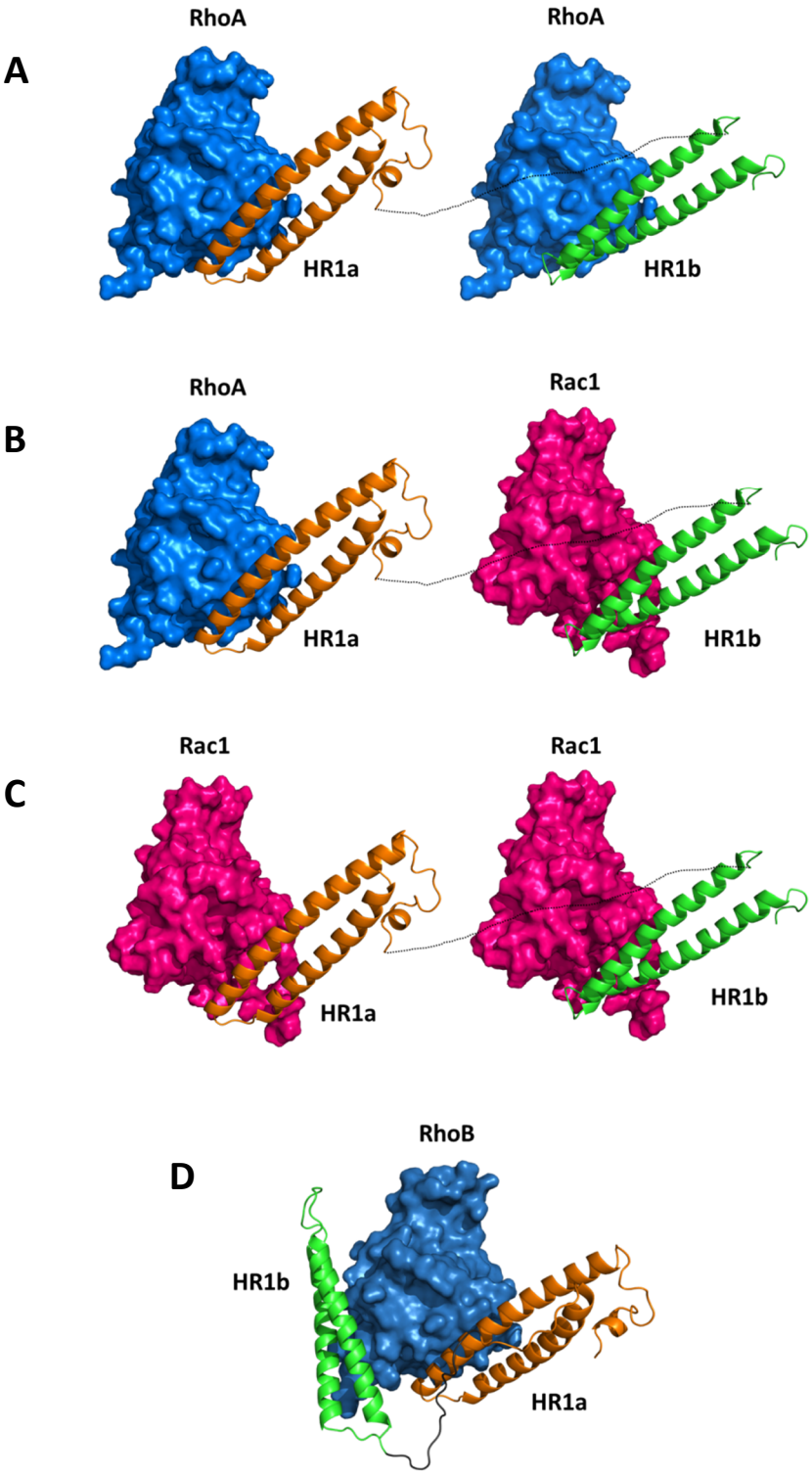


Figure 7.3 Possible supercomplexes between PRK1 HR1 domains and Rho family GTPases. (**A**) One RhoA molecule binds to HR1a while another RhoA molecule binds to HR1b. (**B**) One Rac1 molecule binds to HR1a while another Rac1 molecule binds to HR1b. (**C**) One Rac1 molecule binds to each of HR1a and HR1b. (**D**) Model of PRK1 oligomers and RhoA linked equilibria. All these interactions may occur with HR1 domains of the same protomer or of different protomers.

To analyse these supercomplexes, analytical ultracentrifugation can be used to determine the molecular weight of the complexes formed and as a result determine the stoichiometry of the interacting species. These supercomplexes can then be studied by cryo electron microscopy to obtain structural information on how they assemble. However given the likely polydispersity of these samples, SEC-SAXS may be a better alternative to study these different complexes.

7.2.3.5 Effector handover

RhoA, RhoB, RhoC, Rac1, RhoD and Cdc42 are the six classical Rho GTPases and they all interact with one or more PRKs. Importantly, their cellular localisation differs. Fluorescent studies in cells would therefore be key to understanding if each of these GTPases co-localises with the PRKs to compartments upon particular cellular stimuli. It is also possible that the GTPases are involved in an effector handover mechanism as depicted in Figure 7.1. For example, RhoA activates PRK1 at the plasma membrane and subsequently RhoD binds to PRK1 to take it to early endosomes. PRK1 is then passed on to RhoB in late endosomes where Rac1 can 'recycle' the effector by bringing it back to the plasma membrane. Such antagonistic signalling between RhoB and Rac1 has previously been described (Marcos-Ramiro *et al.*, 2016), however the molecular mechanism and effectors involved in this signalling have not been elucidated. PRK1, a common effector of RhoB and Rac1, may be what is driving this. The PRKs may therefore be involved in a tug-of-war between several Rho GTPases; the GTPase which wins the tug-of-war may depend on the particular signalling pathway.

7.2.4 PRKs on the membrane

The C2-domain of PRKs interacts with membrane lipids and therefore localises the PRKs to the plasma membrane, where they can be activated by GTPases, lipids and PDK1. Biophysical (AUC) and structural studies (NMR) would reveal whether oligomerisation is favoured in the context of the membrane. Detergent bicelles and other membrane mimetics

could be used to immobilise the PRK and then investigate its structure and measure its activity in a kinase assay against some of its many substrates.

7.3 Concluding remarks

The PRKs are an example of how little we know in molecular terms about many signalling proteins. Careful biochemical, biophysical and structural characterisation of any interacting systems is a prerequisite to be able to truly decipher, in molecular terms, how these signalling molecules are activated, regulated and more importantly deregulated in disease. This includes determining the oligomeric state of proteins, the affinity and stoichiometry of interacting domains and the structural characterisation of the complex and in particular of the interaction interface; none of the above can be assumed.

Appendix 1

PRK1 HR1a (1-106):

GPLGSPMASDAVQSEPRSWSLLEQLGLAGADLAAPGVQQQLELERERLRREIRKELKLKEGAENLRRATTDLGRSL
GPVELLLRGSSRRLDLLHQQLQELHAHVVLPDPAAT

PRK1 HR1b (122-199):

GPLGSATNLSRVAGLEKQLAIELKVKQGAENMIQTYSNGSTKDRKLLLTAQQMLQDSKTKIDIIRMQLRRALQAGQ
LENQAAP

PRK1 HR1c (201-297):

GPLGSHMDTQGSPDLGAVELRIEELRHHFRVEHAVAEGAKNVLRLLSAAKAPDRKAVSEAQEKLTESNQKLGLLRE
ALERRLGELPADHPKGRLLREELAAASS

PRK1 HR1ab (1-199):

GPLGSPMASDAVQSEPRSWSLLEQLGLAGADLAAPGVQQQLELERERLRREIRKELKLKEGAENLRRATTDLGRSL
GPVELLLRGSSRRLDLLHQQLQELHAHVVLPDPAATHDGPQSPGAGGPTCSATNLSRVAGLEKQLAIELKVKQGAE
NMIQTYSNGSTKDRKLLLTAQQMLQDSKTKIDIIRMQLRRALQAGQLENQAAP

PRK1 HR1bc (122-297):

GPLGSHMATNLSRVAGLEKQLAIELKVKQGAENMIQTYSNGSTKDRKLLLTAQQMLQDSKTKIDIIRMQLRRALQ
AGQLENQAAPDDTQGSPDLGAVELRIEELRHHFRVEHAVAEGAKNVLRLLSAAKAPDRKAVSEAQEKLTESNQKL
GLLREALERRLGELPADHPKGRLLREELAAASS

PRK1 HR1abc (1-297):

GPLGSHMASDAVQSEPRSWSLLEQLGLAGADLAAPGVQQQLELERERLRREIRKELKLKEGAENLRRATTDLGRSL
GPVELLLRGSSRRLDLLHQQLQELHAHVVLPDPAATHDGPQSPGAGGPTCSATNLSRVAGLEKQLAIELKVKQGAE
NMIQTYSNGSTKDRKLLLTAQQMLQDSKTKIDIIRMQLRRALQAGQLENQAAPDDTQGSPDLGAVELRIEELRHH
FRVEHAVAEGAKNVLRLLSAAKAPDRKAVSEAQEKLTESNQKLGLLREALERRLGELPADHPKGRLLREELAAASS

PRK2 HR1c (207-302):

NAKPVISPLELRMEELRHHFRIEFAVAEGAKNVMKLLGSGKVTDRKALSEAQARFNESSQKLDLLKYSLEQRLNEVP
KNHPKSRIIIEELSLVAAS

PRK2 HR1bc (132-302):

STSNNRLKALQKQLDIELKVKQGAENMIQMYSNGSSKDRKLHGTAQQLLQDSKTKIEVIRMQILQAVQTNELAFD
NAKPVISPLELRMEELRHHFRIEFAVAEGAKNVMKLLGSGKVTDRKALSEAQARFNESSQKLDLLKYSLEQRLNEVP
KNHPKSRIIIEELSLVAAS

PRK2 HR1abc (1-302):

MASNPERGEILLTELQGDSRSLPFSENVSAVQKLDFSDTMVQQKLDDIKDRIKREIRKELKIKEGAENLRKVTTDKKS
LAYVDNILKKSNKKLEELHHKLQELNAHIVVSDPEDITDCPRTPDTPNNDPRCSTSNNRLKALQKQLDIELKVKQGAE
NMIQMYSNGSSKDRKLHGTAQQLLQDSKTKIEVIRMQILQAVQTNELAFDNAKPVISPLELRMEELRHHFRIEFAV
AEGAKNVMKLLGSGKVTDRKALSEAQARFNESSQKLDLLKYSLEQRLNEVPKNHPKSRIIIEELSLVAAS

PRK3 HR1a (1-80):

GPLGSHMEEGAPRQPGPSQWPPEDEKEVIRRAIQKELKIKEGVENLRRVATDRRHLGHVQQLLRSSNRRLEQLHG
ELRELHARILL

PRK3 HR1b (100-171):

GPLGSHMEQLRARHLEALRRQLHVELKVKQGAENMTHTCASGTPKERKLLAAAQQMLRDSQLKVALLRMKISSLE
ASGS

PRK3 HR1c (169-263):

GPLGSHMSGSPEPGPELLAEELQHRLHVEAAVAEGAKNVVKLLSSRRTQDRKALAEAQAQLQESSQKLDLLRLALE
QLLEQLPPAHPLRSRVTRELRAAVPG

PRK3 HR1ab (1-171):

GPLGSHMEEGAPRQPGPSQWPPEDEKEVIRRAIQKELKIKEGVENLRRVATDRRHLGHVQQLLRSSNRRLEQLHG
ELRELHARILLPGPGPGPAEPVASGPRPWAEQLRARHLEALRRQLHVELKVKQGAENMTHTCASGTPKERKLLAAA
QQMLRDSQLKVALLRMKISSLEASGS

PRK3 HR1bc (100-263):

GPLGSHMEQLRARHLEALRRQLHVELKVKQGAENMTHTCASGTPKERKLLAAAQQMLRDSQLKVALLRMKISSLE
ASGSPEPGPELLAEELQHRLHVEAAVAEGAKNVVKLLSSRRTQDRKALAEAQAQLQESSQKLDLLRLALEQLLEQLP
PAHPLRSRVTRELRAAVPG

PRK3 HR1abc (1-263):

GPLGSHMEEGAPRQPGPSQWPPEDEKEVIRRAIQKELKIKEGVENLRRVATDRRHLGHVQQLLRSSNRRLEQLHG
ELRELHARILLPGPGPGPAEPVASGPRPWAEQLRARHLEALRRQLHVELKVKQGAENMTHTCASGTPKERKLLAAA
QQMLRDSQLKVALLRMKISSLEASGSPEPGPELLAEELQHRLHVEAAVAEGAKNVVKLLSSRRTQDRKALAEAQAQ
LQESSQKLDLLRLALEQLLEQLPPAHPLRSRVTRELRAAVPG

RhoA Δ4 F25N Q63L (1-194):

GPLGSMAAIRKKLVIVGDGACGKTCLLIVNSKDQFPEVYVPTVFENYVADIEVDGKQVELALWDTAGLEDYDRLRP
LSYPDTDVILMCFSIDSPDSLENIPEKWTPEVKHFCPNVPIILVGNKKDLRNDEHTRRELAKMKQEPVKPEEGRDMA
NRIGAFGYMECSAKTKDGVREVFEMATRAALQARRGKKKSG

RhoD Δ4 Q75L (1-211):

GPLGSMTAAQAAGEEAPPGVRSVKVVLVGDGGCGKTSLLMVFADGAFPESYTPTVFERYMVNLQVKGKPVHLHI
WDTAGLDDYDRLRPLFYPDASVLLLCFDVTSPNSFDNIFNRWYPEVNHFCKKVPIIVVGCKTDLRKDKSLVNKLRRN
GLEPVTYHRGQEMARSVGAVAYLECSARLHDNVHAVFQEAAEVALSSRGRNFWRRITQGF

Protein sequences. Amino acids from the vector sequence and restriction sites are in bold.

Appendix 2

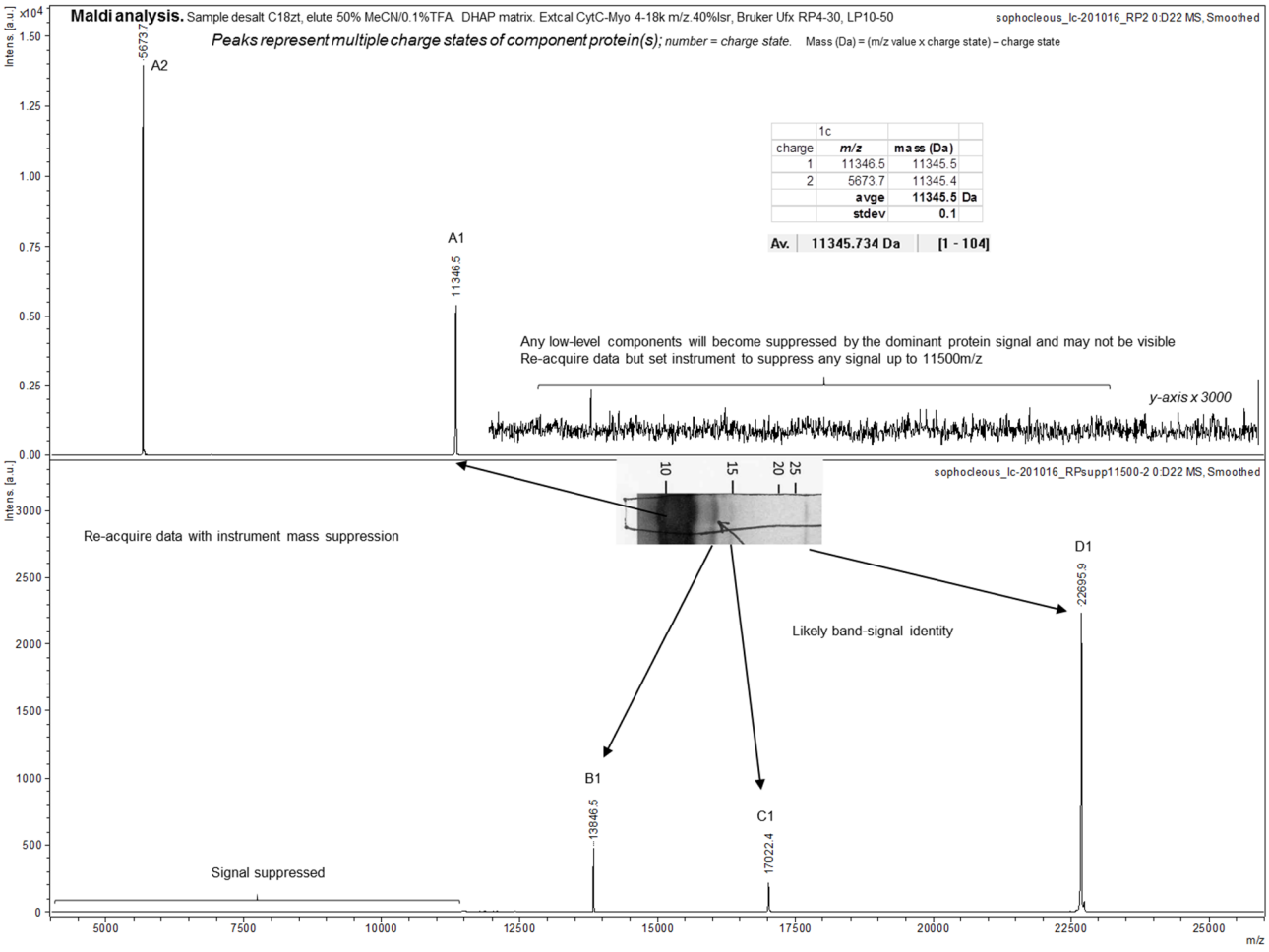


Figure A2.1 PRK1 HR1c MALDI analysis. Predicted MW: 11345.86 Da

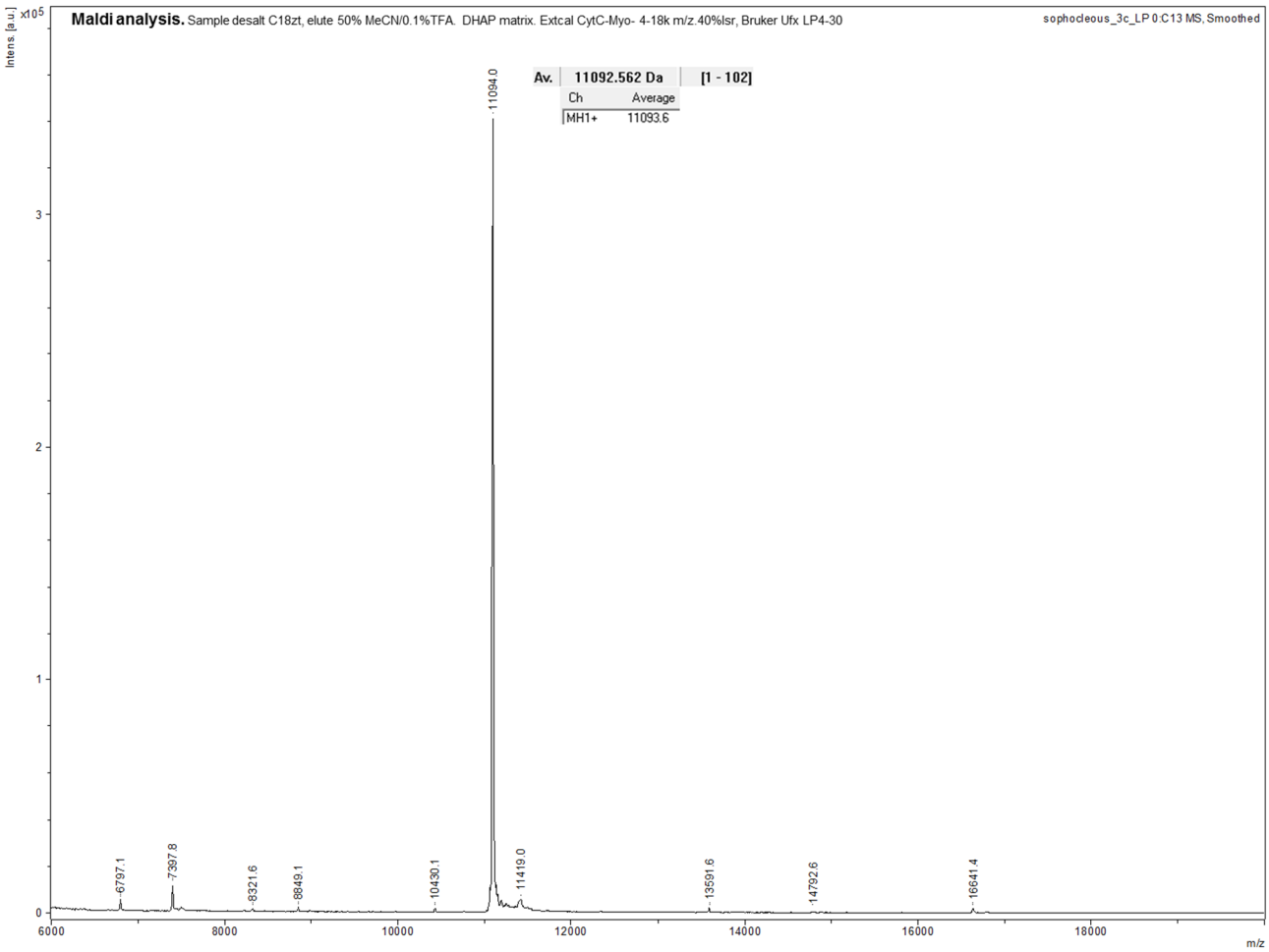


Figure A2.2 PRK3 HR1c MALDI analysis. Predicted MW: 11092.68 Da

Appendix 3

	Scans	Parameters	Concentration 1	Concentration 2	Concentration 3	Concentration 4
PRK1 HR1a	Every 5th out of 999	$\bar{v}$: 0.73239 ρ: 1.0076 η: 0.0161260	C: 100 µM $\frac{f}{f_0}$: 1.34 rmsd: 0.0268 1) Monomer *s*: 0.974 (1.61) MW: 14.9 kDa 59.6% 2) Dimer *s*: 1.64 (2.71) MW: 34.6 kDa 40.4%	C: 200 µM $\frac{f}{f_0}$: 1.29 rmsd: 0.0485 1) Monomer *s*: 1.04 (1.72) MW: 15.6 kDa 41.4% 2) Dimer *s*: 1.77 (2.92) MW: 34.4 kDa 58.6%	C: 400 µM $\frac{f}{f_0}$: 1.36 rmsd: 0.1069 1) Monomer *s*: 1.15 (1.90) MW: 19.4 kDa 26.5% 2) Dimer *s*: 1.79 (2.95) MW: 37.5 kDa 73.5%	
PRK1 HR1b	Every 5th out of 999	$\bar{v}$: 0.73533 ρ: 1.0076 η: 0.0161260	C: 200 µM $\frac{f}{f_0}$: 1.37 rmsd: 0.0284 1) Monomer *s*: 0.737 (1.22) MW: 10.3 kDa 100%	C: 400 µM $\frac{f}{f_0}$: 1.51 rmsd: 0.0402 1) Monomer *s*: 0.726 (1.20) MW: 11.6 kDa 100%	C: 800 µM $\frac{f}{f_0}$: 1.38 rmsd: 0.0349 1) Monomer *s*: 0.729 (1.20) MW: 10.3 kDa 100%	
PRK1 HR1c	Every 5th out of 999	$\bar{v}$: 0.72914 ρ: 1.0076 η: 0.0161260	C: 400 µM $\frac{f}{f_0}$: 1.55 rmsd: 0.0372 1) Monomer *s*: 0.768 (1.27) MW: 12.6 100%	C: 800 µM $\frac{f}{f_0}$: 1.63 rmsd: 0.0137 1) Monomer *s*: 0.773 (1.28) MW: 13.8 100%		
PRK1 HR1ab	Every 5th out of 999	$\bar{v}$: 0.72971 ρ: 1.0078 η: 0.0161487	C: 100 µM $\frac{f}{f_0}$: 1.20 rmsd: 0.0266 1) Monomer *s*: 1.28 (2.12) MW: 18.7 kDa 56.6% 2) Dimer *s*: 2.29 (3.79) MW: 44.7 kDa 43.4%	C: 200 µM $\frac{f}{f_0}$: 1.32 rmsd: 0.0775 1) Monomer *s*: 1.32 (2.18) MW: 22.5 kDa 39.1% 2) Dimer *s*: 2.28 (3.77) MW: 51.0 kDa 60.9%	C: 400 µM $\frac{f}{f_0}$: 1.53 rmsd: 0.2030 1) Monomer *s*: 1.33 (2.20) MW: 28.5 kDa 22.5% 2) Dimer *s*: 2.23 (3.69) MW: 61.9 kDa 77.5%	
PRK1 HR1abc	Every 2nd out of 400	$\bar{v}$: 0.72893 ρ: 1.0078 η: 0.0161487	C: 100 µM $\frac{f}{f_0}$: 1.57 rmsd: 0.0330 1) Monomer *s*: 1.36 (2.25) MW: 30.6 kDa 47.8% 2) Dimer *s*: 2.02 (3.34) MW: 55.5 kDa 23.2%	C: 400 µM $\frac{f}{f_0}$: 1.65 rmsd: 0.2514 1) Monomer *s*: 1.30 (2.15) MW: 31.0 kDa 9.1% 2) Dimer *s*: 1.69 (2.79) MW: 45.8 kDa 18.2%	C: 450 µM $\frac{f}{f_0}$: 1.90 rmsd: 0.2858 1) Monomer *s*: 1.30 (2.15) MW: 38.7 kDa 10.8% 2) Dimer *s*: 1.77 (2.92) MW: 59.9 kDa 15.5%	C: 500 µM $\frac{f}{f_0}$: 1.93 rmsd: 0.3512 1) Monomer *s*: 1.38 (2.28) MW: 43.2 kDa 10.7% 2) Dimer *s*: 1.75 (2.90)

			3) Trimer *s*: 2.74 (4.54) MW: 87.6 kDa 29.0%	3) Trimer *s*: 2.53 (4.19) MW: 83.6 kDa 72.7%	3) Trimer *s*: 2.51 (4.16) MW: 102 kDa 73.7%	MW: 60.9 kDa 10.8% 3) Trimer *s*: 2.52 (4.16) MW: 104 kDa 78.5%
PRK3 HR1a	Every 5th out of 999	$\bar{v}$: 0.72695 ρ: 1.0076 η: 0.0161260	C: 100 µM $\frac{f}{f_0}$: 1.48 rmsd: 0.0137 1) Monomer *s*: 0.782 (1.29) MW: 12.0 kDa 89.9% 2) Dimer *s*: 1.49 (2.46) MW: 31.8 kDa 10.1%	C: 200 µM $\frac{f}{f_0}$: 1.60 rmsd: 0.0302 1) Monomer *s*: 0.713 (1.18) MW: 11.7 kDa 78.1% 2) Dimer *s*: 1.15 (1.90) MW: 24.0 kDa 21.9%	C: 400 µM $\frac{f}{f_0}$: 1.53 rmsd: 0.0409 1) Monomer *s*: 0.729 (1.20) MW: 11.3 kDa 71.9% 2) Dimer *s*: 1.15 (1.89) MW: 22.3 kDa 28.1%	
PRK3 HR1b	Every 5th out of 999	$\bar{v}$: 0.73257 ρ: 1.0221 η: 0.0166578	C: 200 µM $\frac{f}{f_0}$: 1.45 rmsd: 0.0224 1) Monomer *s*: 0.716 (1.27) MW: 11.8 kDa 90.3% 2) Dimer *s*: 1.19 (2.12) MW: 25.2 kDa 9.7%	C: 400 µM $\frac{f}{f_0}$: 1.50 rmsd: 0.0313 1) Monomer *s*: 0.727 (1.29) MW: 12.7 kDa 85.1% 2) Dimer *s*: 1.26 (2.24) MW: 28.8 kDa 14.9%		
PRK3 HR1c	Every 3rd out of 600	$\bar{v}$: 0.73364 ρ: 1.0076 η: 0.0161260	C: 100 µM $\frac{f}{f_0}$: 1.70 rmsd: 0.0181 1) Monomer *s*: 0.666 (1.10) MW: 12.7 kDa 49.1% 2) Dimer *s*: 0.949 (1.57) MW: 20.5 kDa 44.7% 3) Tetramer *s*: 1.68 (2.78) MW: 48.6 kDa 6.2%	C: 200 µM $\frac{f}{f_0}$: 1.71 rmsd: 0.0141 1) Monomer *s*: 0.606 (1.00) MW: 10.9 kDa 36.4% 2) Dimer *s*: 1.02 (1.68) MW: 23.2 kDa 56.8% 3) Tetramer *s*: 1.63 (2.69) MW: 46.9 kDa 6.8%	C: 400 µM $\frac{f}{f_0}$: 1.71 rmsd: 0.0279 1) Monomer *s*: 0.731 (1.21) MW: 14.2 kDa 34.7% 2) Dimer *s*: 1.10 (1.82) MW: 25.9 kDa 49.9% 3) Tetramer *s*: 1.51 (2.49) MW: 41.4 kDa 15.4%	
PRK3 HR1ab	Every 5th out of 999	$\bar{v}$: 0.72894 ρ: 1.0221 η: 0.0166578	C: 50 µM $\frac{f}{f_0}$: 1.31 rmsd: 0.0207 1) Monomer *s*: 1.08 (1.92) MW: 18.5 kDa 91.4% 2) Dimer *s*: 1.55 (2.76)			

			MW: 33.0 kDa 8.6%			
PRK3 HR1abc	Every one out of 200	$\bar{v}$: 0.73175 ρ: 1.0221 η: 0.0166578	C: 50 µM $\frac{f}{f_0}$: 1.07 rmsd: 0.0363 1) Monomer *s*: 1.42 (2.53) MW: 20.8 kDa 13.5% 2) Dimer *s*: 2.10 (3.74) MW: 37.3 kDa 22.4% 3) Trimer *s*: 2.78 (4.94) MW: 58.1 kDa 55.8% 4) Larger oligomer *s*: 4.04 (7.18) MW: 98.7 kDa 8.3%	C: 100 µM $\frac{f}{f_0}$: 1.14 rmsd: 0.0443 1) Monomer *s*: 1.51 (2.68) MW: 24.9 kDa 12.9% 2) Dimer *s*: 2.34 (4.15) MW: 48.8 kDa 19.1% 3) Trimer *s*: 2.89 (5.14) MW: 66.2 kDa 49.5% 4) Larger oligomer *s*: 4.13 (7.35) MW: 113 kDa 18.5%	C: 200 µM $\frac{f}{f_0}$: 1.07 rmsd: 0.0655 1) Monomer *s*: 1.47 (2.52) MW: 21.1 kDa 5.1% 2) Dimer *s*: 2.23 (3.81) MW: 38.8 kDa 14.1% 3) Trimer *s*: 3.01 (5.13) MW: 60.4 kDa 54.5% 4) Larger oligomer *s*: 4.51 (7.70) MW: 111 kDa 26.3%	
RhoA Δ4 F25N Q63L + PRK1 HR1a	Every one out of 200	$\bar{v}$: 0.7298 ρ: 1.0082 η: 0.01618	C: 150 µM RhoA Δ4 F25N Q63L $\frac{f}{f_0}$: 1.11 rmsd: 0.021 1) RhoA *s*: 1.53 (2.53) MW: 22.1 kDa 100%	C: 150 µM RhoA Δ4 F25N Q63L + 150 µM PRK1 HR1a $\frac{f}{f_0}$: 1.31 rmsd: 0.041 1) RhoA *s*: 1.54 (2.53) MW: 21.9 kDa 22.8% 2) RhoA + HR1a complex *s*: 2.64 (4.36) MW: 49.6 kDa 77.2%		

Table A3.1 Summary of AUC results. All experiments were SV experiments carried out at 40000 rpm at 4 ^{o}C. Sedimentation was monitored by interference. A total of 200 scans, recorded with an interval of 1 min 30 sec, were analysed for each different protein concentration. The partial specific volume, buffer density and buffer viscosity were estimated using the program SEDNTERP. The frictional ratio and rmsd of each fit, as well as the *s*-value ($s_{20,w}$ in brackets), % abundance and corresponding MW of each species detected are listed. Artefacts and the buffer mismatch peak at *s*=0 were excluded from the percentage calculations.

Appendix 4

	H	N	C_α	C_β	
194 Gly			43.64		$H^{\alpha a}$ 3.88, $H^{\alpha b}$ 3.88
195 Pro			63.18	32.36	H^{α} 4.40, $H^{\beta a}$ 1.87, $H^{\beta b}$ 2.24, $H^{\gamma a}$ 1.94, $H^{\gamma b}$ 1.94, $H^{\delta a}$ 3.50, $H^{\delta b}$ 3.50, C^{γ} 27.17, C^{δ} 49.70
196 Leu	8.46	121.96	55.44	42.25	H^{α} 4.28, $H^{\beta a}$ 1.54, $H^{\beta b}$ 1.59, H^{γ} 1.59, $H^{\delta a*}$ 0.81, $H^{\delta b*}$ 0.86, C^{γ} 27.07, $C^{\delta a}$ 23.62, $C^{\delta b}$ 24.96
197 Gly	8.33	109.57	45.17		$H^{\alpha a}$ 3.90, $H^{\alpha b}$ 3.90
198 Ser	8.19	116.53	56.38	63.50	H^{α} 4.72, $H^{\beta a}$ 3.78, $H^{\beta b}$ 3.83
199 His			56.80	30.64	H^{α} 4.53, $H^{\beta a}$ 3.04, $H^{\beta b}$ 3.04, $H^{\delta 2}$ 6.93, $H^{\varepsilon 1}$ 7.84, $C^{\delta 2}$ 119.78, $C^{\varepsilon 1}$ 138.34
200 Met	8.07	120.66	55.51	32.91	H^{α} 4.36, $H^{\beta a}$ 1.84, $H^{\beta b}$ 1.97, $H^{\gamma a}$ 2.30, $H^{\gamma b}$ 2.40, $H^{\varepsilon *}$ 1.99, C^{γ} 31.92, C^{ε} 16.93
201 Asp	8.34	121.35	54.59	41.25	H^{α} 4.59, $H^{\beta a}$ 2.61, $H^{\beta b}$ 2.66
202 Thr	8.04	113.60	61.98	69.63	H^{α} 4.25, H^{β} 4.23, $H^{\gamma 2*}$ 1.13, $C^{\gamma 2}$ 21.69
203 Gln	8.33	121.97	56.18	29.44	H^{α} 4.27, $H^{\beta a}$ 1.97, $H^{\beta b}$ 2.08, $H^{\gamma a}$ 2.31, $H^{\gamma b}$ 2.31, C^{γ} 33.98
203 Gln					$N^{\varepsilon 2}$ 112.27, $H^{\varepsilon 2a}$ 6.77, $H^{\varepsilon 2b}$ 7.51
204 Gly	8.34	109.62	45.30		$H^{\alpha a}$ 3.90, $H^{\alpha b}$ 3.92
205 Ser	8.15	115.24			H^{α} 4.37, $H^{\beta b}$ 3.78
205 Ser	7.91	114.99	57.11	64.00	$H^{\beta a}$ 3.74
206 Pro			63.41	32.18	H^{α} 4.38, $H^{\beta a}$ 1.81, $H^{\beta b}$ 2.21, $H^{\gamma a}$ 1.95, $H^{\gamma b}$ 1.95, $H^{\delta a}$ 3.68, $H^{\delta b}$ 3.73, C^{γ} 27.28, C^{δ} 50.78
207 Asp	8.30	120.12	53.92	41.17	H^{α} 4.56, $H^{\beta a}$ 2.58, $H^{\beta b}$ 2.74
208 Leu	8.36	123.98	56.50	41.84	H^{α} 4.16, $H^{\beta a}$ 1.57, $H^{\beta b}$ 1.68, H^{γ} 1.67, $H^{\delta a*}$ 0.85, $H^{\delta b*}$ 0.90, C^{γ} 26.99, $C^{\delta a}$ 23.81, $C^{\delta b}$ 25.24
209 Gly	8.43	107.73	46.84		$H^{\alpha a}$ 3.86, $H^{\alpha b}$ 3.86
210 Ala	7.91	123.51	54.62	18.52	H^{α} 4.14, $H^{\beta *}$ 1.43
211 Val	7.83	119.58	66.14	32.00	H^{α} 3.58, H^{β} 2.09, $H^{\gamma a*}$ 0.83, $H^{\gamma b*}$ 0.98, $C^{\gamma a}$ 21.86, $C^{\gamma b}$ 22.46
212 Glu	8.39	119.21	60.34	29.77	H^{α} 3.72, $H^{\beta a}$ 1.96, $H^{\beta b}$ 2.12, $H^{\gamma a}$ 2.02, $H^{\gamma b}$ 2.36, C^{γ} 37.61
213 Leu	7.84	119.57	57.81	41.94	H^{α} 4.08, $H^{\beta a}$ 1.57, $H^{\beta b}$ 1.73, H^{γ} 1.68, $H^{\delta a*}$ 0.83, $H^{\delta b*}$ 0.84, C^{γ} 27.04, $C^{\delta a}$ 23.80, $C^{\delta b}$ 24.96
214 Arg	7.87	119.50	57.97	30.04	H^{α} 4.17, $H^{\beta a}$ 1.94, $H^{\beta b}$ 1.94, $H^{\gamma a}$ 1.62, $H^{\gamma b}$ 1.79, $H^{\delta a}$ 3.14, $H^{\delta b}$ 3.31, C^{γ} 26.85, C^{δ} 43.18
215 Ile	8.38	119.12	66.13	37.74	H^{α} 3.48, H^{β} 1.90, $H^{\gamma 1a}$ 0.71, $H^{\gamma 1b}$ 1.70, $H^{\gamma 2*}$ 0.72, $H^{\delta 1*}$ 0.61, $C^{\gamma 1}$ 30.90, $C^{\gamma 2}$ 17.60, $C^{\delta 1}$ 14.46
216 Glu	8.04	118.99	59.46	29.17	H^{α} 4.00, $H^{\beta a}$ 2.08, $H^{\beta b}$ 2.08, $H^{\gamma a}$ 2.14, $H^{\gamma b}$ 2.34, C^{γ} 36.42
217 Glu	7.86	120.37	59.50	30.08	H^{α} 4.07, $H^{\beta a}$ 1.87, $H^{\beta b}$ 2.21, $H^{\gamma a}$ 2.14, $H^{\gamma b}$ 2.38, C^{γ} 36.52
218 Leu	8.12	119.36	58.16	43.31	H^{α} 4.08, $H^{\beta a}$ 1.28, $H^{\beta b}$ 2.05, H^{γ} 1.80, $H^{\delta a*}$ 0.76, $H^{\delta b*}$ 0.90, C^{γ} 27.37, $C^{\delta a}$ 26.00, $C^{\delta b}$ 23.48
219 Arg	8.94	119.00	60.16	30.90	H^{α} 3.85, $H^{\beta a}$ 1.80, $H^{\beta b}$ 2.02, $H^{\gamma a}$ 1.49, $H^{\gamma b}$ 1.77, $H^{\delta a}$ 2.96, $H^{\delta b}$ 3.10, C^{γ} 27.72, C^{δ} 44.17
220 His	7.86	118.58	60.28	30.06	H^{α} 4.31, $H^{\beta a}$ 3.11, $H^{\beta b}$ 3.27, $H^{\delta 2}$ 6.50, $H^{\varepsilon 1}$ 7.81, $C^{\delta 2}$ 119.49, $C^{\varepsilon 1}$ 138.84
221 His	8.07	117.00	59.52	31.07	H^{α} 4.22, $H^{\beta a}$ 3.16, $H^{\beta b}$ 3.20, $H^{\delta 2}$ 7.02, $H^{\varepsilon 1}$ 7.63, $C^{\delta 2}$ 119.59, $C^{\varepsilon 1}$ 138.38
222 Phe	9.05	120.79	62.05	39.65	H^{α} 3.98, $H^{\beta a}$ 2.98, $H^{\beta b}$ 3.32, $H^{\delta *}$ 7.02, $H^{\varepsilon *}$ 7.12, H^{ζ} 7.16, $C^{\delta *}$ 131.64, $C^{\varepsilon *}$ 131.25, C^{ζ} 129.59
223 Arg	8.26	118.97	59.98	30.17	H^{α} 3.88, $H^{\beta a}$ 1.87, $H^{\beta b}$ 2.02, $H^{\gamma a}$ 1.68, $H^{\gamma b}$ 1.95, $H^{\delta a}$ 3.18, $H^{\delta b}$ 3.19, C^{γ} 27.94, C^{δ} 43.69
224 Val	8.05	120.28	66.79	32.16	H^{α} 3.50, H^{β} 2.01, $H^{\gamma a*}$ 0.67, $H^{\gamma b*}$ 0.81, $C^{\gamma a}$ 22.58, $C^{\gamma b}$ 21.23
225 Glu	8.23	119.83	58.36	28.91	H^{α} 3.91, $H^{\beta a}$ 1.80, $H^{\beta b}$ 1.94, $H^{\gamma a}$ 2.23, $H^{\gamma b}$ 2.38, C^{γ} 36.48
226 His	8.76	118.92	60.73	29.95	H^{α} 3.89, $H^{\beta a}$ 2.66, $H^{\beta b}$ 2.90, $H^{\delta 2}$ 6.75, $H^{\varepsilon 1}$ 7.77, $C^{\delta 2}$ 119.47, $C^{\varepsilon 1}$ 138.50
227 Ala	7.43	120.19	54.95	18.03	H^{α} 4.11, $H^{\beta *}$ 1.45
228 Val	7.89	123.02	66.66	31.19	H^{α} 3.49, H^{β} 2.38, $H^{\gamma a*}$ 0.86, $H^{\gamma b*}$ 0.97, $C^{\gamma a}$ 21.43, $C^{\gamma b}$ 22.45
229 Ala	8.38	125.12	56.03	17.66	H^{α} 3.80, $H^{\beta *}$ 1.41
230 Glu	8.35	117.09	59.24	29.69	H^{α} 3.78, $H^{\beta a}$ 1.90, $H^{\beta b}$ 1.92, $H^{\gamma a}$ 2.03, $H^{\gamma b}$ 2.05, C^{γ} 36.32
231 Gly	8.03	107.11	47.19		$H^{\alpha a}$ 3.86, $H^{\alpha b}$ 3.86
232 Ala	8.22	125.21	55.02	17.87	H^{α} 4.13, $H^{\beta *}$ 1.43

233 Lys	8.54	119.05	60.13	32.59	H^{α} 3.83, $H^{\beta a}$ 1.83, $H^{\beta b}$ 1.88, $H^{\gamma a}$ 1.35, $H^{\gamma b}$ 1.60, $H^{\delta a}$ 1.61, $H^{\delta b}$ 1.62, $H^{\varepsilon a}$ 2.76, $H^{\varepsilon b}$ 2.86, C^{γ} 25.99, C^{δ} 29.75, C^{ε} 42.07
234 Asn					$N^{\delta 2}$ 113.44
234 Asn	7.96	118.54	56.46	38.37	H^{α} 4.42, $H^{\beta a}$ 2.83, $H^{\beta b}$ 2.93, $H^{\delta 2a}$ 7.03, $H^{\delta 2b}$ 7.58
235 Val	7.99	120.51	67.19	32.03	H^{α} 3.54, H^{β} 2.09, $H^{\gamma a*}$ 0.87, $H^{\gamma b*}$ 1.03, $C^{\gamma a}$ 21.32, $C^{\gamma b}$ 23.59
236 Leu	8.04	119.11	58.38	41.92	H^{α} 3.93, $H^{\beta a}$ 1.58, $H^{\beta b}$ 1.70, H^{γ} 1.58, $H^{\delta a*}$ 0.84, $H^{\delta b*}$ 0.89, C^{γ} 27.16, $C^{\delta a}$ 25.08, $C^{\delta b}$ 25.93
237 Arg	8.13	120.42	59.63	30.01	H^{α} 3.98, $H^{\beta a}$ 1.93, $H^{\beta b}$ 1.93, $H^{\gamma a}$ 1.52, $H^{\gamma b}$ 1.70, $H^{\delta a}$ 3.15, $H^{\delta b}$ 3.15, C^{γ} 27.51, C^{δ} 43.36
238 Leu	7.77	120.20	57.89	42.26	H^{α} 4.05, $H^{\beta a}$ 1.46, $H^{\beta b}$ 1.87, H^{γ} 1.70, $H^{\delta a*}$ 0.80, $H^{\delta b*}$ 0.81, C^{γ} 27.24, $C^{\delta a}$ 25.82, $C^{\delta b}$ 23.31
239 Leu	8.59	119.57	57.89	42.75	H^{α} 4.03, $H^{\beta a}$ 1.40, $H^{\beta b}$ 1.84, H^{γ} 1.85, $H^{\delta a*}$ 0.68, $H^{\delta b*}$ 0.87, C^{γ} 26.87, $C^{\delta a}$ 27.16, $C^{\delta b}$ 22.92
240 Ser	8.06	112.06	60.91	63.29	H^{α} 4.13, $H^{\beta a}$ 3.89, $H^{\beta b}$ 3.94
241 Ala	7.22	121.80	52.00	19.18	H^{α} 4.31, $H^{\beta *}$ 1.42
242 Ala	6.98	121.24	51.92	19.53	H^{α} 4.04, $H^{\beta *}$ 1.21
243 Lys	8.28	120.49	58.33	32.19	H^{α} 3.94, $H^{\beta a}$ 1.72, $H^{\beta b}$ 1.72, $H^{\gamma a}$ 1.39, $H^{\gamma b}$ 1.43, $H^{\delta a}$ 1.63, $H^{\delta b}$ 1.63, $H^{\varepsilon a}$ 2.93, $H^{\varepsilon b}$ 2.93, C^{γ} 24.95, C^{δ} 29.01, C^{ε} 42.08
244 Ala	8.06	121.22	49.39	18.51	H^{α} 4.69, $H^{\beta *}$ 1.17
245 Pro			63.01	32.49	H^{α} 4.26, $H^{\beta a}$ 1.62, $H^{\beta b}$ 2.15, $H^{\gamma a}$ 1.90, $H^{\gamma b}$ 1.94, $H^{\delta a}$ 3.48, $H^{\delta b}$ 3.48, C^{γ} 27.19, C^{δ} 50.49
246 Asp	8.74	122.56	53.26	41.34	H^{α} 4.53, $H^{\beta a}$ 2.57, $H^{\beta b}$ 2.84
247 Arg	8.54	125.46	59.39	29.85	H^{α} 3.81, $H^{\beta a}$ 1.82, $H^{\beta b}$ 1.82, $H^{\gamma a}$ 1.64, $H^{\gamma b}$ 1.72, $H^{\delta a}$ 3.15, $H^{\delta b}$ 3.15, C^{γ} 27.32, C^{δ} 43.24
248 Lys	8.14	120.67	59.21	32.14	H^{α} 4.05, $H^{\beta a}$ 1.83, $H^{\beta b}$ 1.88, $H^{\gamma a}$ 1.30, $H^{\gamma b}$ 1.40, $H^{\delta a}$ 1.63, $H^{\delta b}$ 1.63, $H^{\varepsilon a}$ 2.92, $H^{\varepsilon b}$ 2.92, C^{γ} 24.85, C^{δ} 29.11, C^{ε} 42.05
249 Ala	7.65	122.68	54.78	18.40	H^{α} 4.09, $H^{\beta *}$ 1.43
250 Val	8.30	118.98	66.99	32.14	H^{α} 3.38, H^{β} 1.98, $H^{\gamma a*}$ 0.80, $H^{\gamma b*}$ 0.91, $C^{\gamma a}$ 21.13, $C^{\gamma b}$ 23.52
251 Ser	8.03	114.69	62.17	62.75	H^{α} 4.16, $H^{\beta a}$ 3.95, $H^{\beta b}$ 3.95
252 Glu	8.17	121.18	59.53	29.71	H^{α} 4.00, $H^{\beta a}$ 1.93, $H^{\beta b}$ 2.04, $H^{\gamma a}$ 2.15, $H^{\gamma b}$ 2.42, C^{γ} 36.78
253 Ala	7.88	122.45	55.10	18.15	H^{α} 4.13, $H^{\beta *}$ 1.45
254 Gln					$N^{\varepsilon 2}$ 111.29
254 Gln	8.74	118.78	59.12	28.41	H^{α} 3.96, $H^{\beta a}$ 2.10, $H^{\beta b}$ 2.26, $H^{\gamma a}$ 2.31, $H^{\gamma b}$ 2.39, $H^{\varepsilon 2a}$ 6.69, $H^{\varepsilon 2b}$ 7.15, C^{γ} 34.34
255 Glu	8.23	120.72	59.69	29.41	H^{α} 4.01, $H^{\beta a}$ 2.06, $H^{\beta b}$ 2.13, $H^{\gamma a}$ 2.34, $H^{\gamma b}$ 2.38, C^{γ} 36.44
256 Lys	7.83	119.20	58.83	32.08	H^{α} 4.04, $H^{\beta a}$ 1.79, $H^{\beta b}$ 1.91, $H^{\gamma a}$ 1.44, $H^{\gamma b}$ 1.48, $H^{\delta a}$ 1.61, $H^{\delta b}$ 1.64, $H^{\varepsilon a}$ 2.88, $H^{\varepsilon b}$ 2.88, C^{γ} 24.93, C^{δ} 28.67, C^{ε} 42.16
257 Leu	8.03	121.18	58.23	41.89	H^{α} 3.94, $H^{\beta a}$ 1.25, $H^{\beta b}$ 2.06, H^{γ} 1.51, $H^{\delta a*}$ 0.82, $H^{\delta b*}$ 0.94, C^{γ} 27.22, $C^{\delta a}$ 27.10, $C^{\delta b}$ 24.00
258 Thr	8.58	116.68	67.00	68.86	H^{α} 3.84, H^{β} 4.24, $H^{\gamma 2*}$ 1.19, $C^{\gamma 2}$ 21.85
259 Glu	8.14	121.27			H^{α} 4.01, $H^{\beta a}$ 2.04, $H^{\gamma b}$ 2.34, C^{γ} 36.45
259 Glu			59.54	29.76	$H^{\beta b}$ 2.10, $H^{\gamma a}$ 2.33
260 Ser	8.03	115.69	62.00	63.63	H^{α} 4.29, $H^{\beta a}$ 3.77, $H^{\beta b}$ 3.85
261 Asn	8.78	118.18	56.04	37.73	H^{α} 4.51, $H^{\beta a}$ 2.67, $H^{\beta b}$ 2.89, $H^{\delta 2a}$ 6.41, $H^{\delta 2b}$ 7.33
261 Asn					$N^{\delta 2}$ 108.55
262 Gln	8.09	121.29	59.17	28.13	$N^{\varepsilon 2}$ 110.65, H^{α} 4.11, $H^{\beta a}$ 2.17, $H^{\beta b}$ 2.23, $H^{\gamma a}$ 2.33, $H^{\gamma b}$ 2.51, $H^{\varepsilon 2a}$ 6.76, $H^{\varepsilon 2b}$ 7.32, C^{γ} 34.07
263 Lys	7.90	121.26	60.18	32.68	H^{α} 3.94, $H^{\beta a}$ 1.75, $H^{\beta b}$ 1.98, $H^{\gamma a}$ 1.35, $H^{\gamma b}$ 1.61, $H^{\delta a}$ 1.61, $H^{\delta b}$ 1.61, $H^{\varepsilon a}$ 2.83, $H^{\varepsilon b}$ 2.83, C^{γ} 25.79, C^{δ} 29.43, C^{ε} 42.12
264 Leu	9.12	120.38	58.38	42.31	H^{α} 3.95, $H^{\beta a}$ 1.56, $H^{\beta b}$ 1.83, H^{γ} 1.57, $H^{\delta a*}$ 0.46, $H^{\delta b*}$ 0.56, C^{γ} 26.76, $C^{\delta a}$ 24.98, $C^{\delta b}$ 25.49
265 Gly	8.01	105.99	47.44		$H^{\alpha a}$ 3.87, $H^{\alpha b}$ 3.97
266 Leu	7.47	122.10	57.97	42.08	H^{α} 4.17, $H^{\beta a}$ 1.53, $H^{\beta b}$ 1.91, H^{γ} 1.78, $H^{\delta a*}$ 0.85, $H^{\delta b*}$ 0.87, C^{γ} 26.98, $C^{\delta a}$ 23.07, $C^{\delta b}$ 25.61
267 Leu	8.28	121.79	58.06	43.71	H^{α} 4.10, $H^{\beta a}$ 1.37, $H^{\beta b}$ 2.23, H^{γ} 1.83, $H^{\delta a*}$ 0.80, $H^{\delta b*}$ 0.95, C^{γ} 27.30, $C^{\delta a}$ 26.02, $C^{\delta b}$ 23.62
268 Arg	8.85	119.52	60.28	30.14	H^{α} 3.68, $H^{\beta a}$ 2.04, $H^{\beta b}$ 2.04, $H^{\gamma a}$ 1.52, $H^{\gamma b}$ 1.58, $H^{\delta a}$ 3.21, $H^{\delta b}$ 3.24, H^{ε} 7.14, C^{γ} 28.28, C^{δ} 43.40
269 Glu	7.81	117.72	59.53	29.61	H^{α} 4.00, $H^{\beta a}$ 2.05, $H^{\beta b}$ 2.10, $H^{\gamma a}$ 2.18, $H^{\gamma b}$ 2.41, C^{γ} 36.57
270 Ala	7.80	121.58	55.25	18.83	H^{α} 4.04, $H^{\beta *}$ 1.53
271 Leu	8.49	119.39	58.49	42.41	H^{α} 3.78, $H^{\beta a}$ 1.29, $H^{\beta b}$ 1.84, H^{γ} 1.52, $H^{\delta a*}$ 0.52, $H^{\delta b*}$ 0.75, C^{γ} 27.81, $C^{\delta a}$ 24.79, $C^{\delta b}$ 25.80

272 Glu	8.55	118.30	60.21	29.43	H^{α} 3.70, $H^{\beta a}$ 1.92, $H^{\beta b}$ 2.08, $H^{\gamma a}$ 2.06, $H^{\gamma b}$ 2.37, C^{γ} 37.31
273 Arg	7.92	119.13	59.44	29.95	H^{α} 3.99, $H^{\beta a}$ 1.82, $H^{\beta b}$ 1.93, $H^{\gamma a}$ 1.53, $H^{\gamma b}$ 1.73, $H^{\delta a}$ 3.14, $H^{\delta b}$ 3.14, C^{γ} 27.62, C^{δ} 43.43
274 Arg	7.86	117.53	57.77	30.45	H^{α} 4.10, $H^{\beta a}$ 1.63, $H^{\beta b}$ 1.87, $H^{\gamma a}$ 1.44, $H^{\gamma b}$ 1.88, $H^{\delta a}$ 3.03, $H^{\delta b}$ 3.13, C^{γ} 26.58, C^{δ} 43.01
275 Leu	8.49	119.01	58.04	41.60	H^{α} 3.78, $H^{\beta a}$ 1.51, $H^{\beta b}$ 1.66, H^{γ} 1.63, $H^{\delta a*}$ 0.75, $H^{\delta b*}$ 0.75, C^{γ} 27.13, $C^{\delta a}$ 24.94, $C^{\delta b}$ 24.95
276 Gly	7.52	103.95	45.99		$H^{\alpha a}$ 3.79, $H^{\alpha b}$ 3.97
277 Glu	7.47	118.09	56.45	30.54	H^{α} 4.18, $H^{\beta a}$ 2.07, $H^{\beta b}$ 2.18, $H^{\gamma a}$ 2.14, $H^{\gamma b}$ 2.38, C^{γ} 36.78
278 Leu	7.16	118.99	52.41	41.84	H^{α} 4.48, $H^{\beta a}$ 1.12, $H^{\beta b}$ 1.71, H^{γ} 1.60, $H^{\delta a*}$ 0.33, $H^{\delta b*}$ 0.38, C^{γ} 25.84, $C^{\delta a}$ 22.27, $C^{\delta b}$ 25.25
279 Pro			62.84	32.49	H^{α} 4.45, $H^{\beta a}$ 2.01, $H^{\beta b}$ 2.44, $H^{\gamma a}$ 2.16, $H^{\gamma b}$ 2.20, $H^{\delta a}$ 3.46, $H^{\delta b}$ 3.90, C^{γ} 27.90, C^{δ} 50.55
280 Ala	8.76	125.10	54.73	18.55	H^{α} 4.01, $H^{\beta *}$ 1.36
281 Asp	8.30	113.40	53.17	39.87	H^{α} 4.48, $H^{\beta a}$ 2.60, $H^{\beta b}$ 2.75
282 His	7.69	122.95	56.67	32.88	H^{α} 4.48, $H^{\beta a}$ 2.91, $H^{\beta b}$ 3.14, $H^{\delta 2}$ 7.07, $H^{\varepsilon 1}$ 7.62, $C^{\delta 2}$ 118.29, $C^{\varepsilon 1}$ 138.93
283 Pro			65.53	32.40	H^{α} 4.22, $H^{\beta a}$ 1.87, $H^{\beta b}$ 2.28, $H^{\gamma a}$ 1.81, $H^{\gamma b}$ 1.86, $H^{\delta a}$ 2.62, $H^{\delta b}$ 3.64, C^{γ} 27.50, C^{δ} 50.90
284 Lys	10.08	119.80	57.90	32.29	H^{α} 4.24, $H^{\beta a}$ 1.81, $H^{\beta b}$ 1.97, $H^{\gamma a}$ 1.49, $H^{\gamma b}$ 1.61, $H^{\delta a}$ 1.57, $H^{\delta b}$ 1.71, $H^{\varepsilon a}$ 2.94, $H^{\varepsilon b}$ 2.94, C^{γ} 25.56, C^{δ} 28.91, C^{ε} 41.74
285 Gly	8.22	107.21	47.54		$H^{\alpha a}$ 3.55, $H^{\alpha b}$ 4.01
286 Arg	7.84	120.38	59.45	29.97	H^{α} 3.98, $H^{\beta a}$ 1.81, $H^{\beta b}$ 1.81, $H^{\gamma a}$ 1.52, $H^{\gamma b}$ 1.64, $H^{\delta a}$ 3.13, $H^{\delta b}$ 3.13, C^{γ} 27.28, C^{δ} 43.45
287 Leu	7.39	118.63	57.69	42.01	H^{α} 4.05, $H^{\beta a}$ 1.57, $H^{\beta b}$ 1.70, H^{γ} 1.62, $H^{\delta a*}$ 0.82, $H^{\delta b*}$ 0.86, C^{γ} 27.12, $C^{\delta a}$ 24.09, $C^{\delta b}$ 24.89
288 Leu	8.06	119.60	57.40	41.82	H^{α} 4.05, $H^{\beta a}$ 1.50, $H^{\beta b}$ 1.70, H^{γ} 1.63, $H^{\delta a*}$ 0.75, $H^{\delta b*}$ 0.75, C^{γ} 27.39, $C^{\delta a}$ 24.80
289 Arg	8.17	118.20	60.24	30.18	H^{α} 3.77, $H^{\beta a}$ 1.84, $H^{\beta b}$ 1.85, $H^{\gamma a}$ 1.51, $H^{\gamma b}$ 1.78, $H^{\delta a}$ 3.06, $H^{\delta b}$ 3.15, C^{γ} 29.08, C^{δ} 43.31
290 Glu	7.81	118.90	59.33	29.63	H^{α} 3.99, $H^{\beta a}$ 2.04, $H^{\beta b}$ 2.10, $H^{\gamma a}$ 2.18, $H^{\gamma b}$ 2.39, C^{γ} 36.57
291 Glu	8.05	121.19	59.46	29.89	H^{α} 3.99, $H^{\beta a}$ 1.97, $H^{\beta b}$ 2.08, $H^{\gamma a}$ 2.06, $H^{\gamma b}$ 2.35, C^{γ} 36.58
292 Leu	8.27	118.84	57.28	41.61	H^{α} 3.88, $H^{\beta a}$ 1.35, $H^{\beta b}$ 1.65, H^{γ} 1.53, $H^{\delta a*}$ 0.36, $H^{\delta b*}$ 0.56, C^{γ} 27.01, $C^{\delta a}$ 22.85, $C^{\delta b}$ 25.42
293 Ala	7.67	121.25	54.09	18.44	H^{α} 4.13, $H^{\beta *}$ 1.40
294 Ala	7.64	120.19	53.39	18.56	H^{α} 4.17, $H^{\beta *}$ 1.40
295 Ala	7.61	120.81	52.35	18.98	H^{α} 4.32, $H^{\beta *}$ 1.34
296 Ser	7.91	114.83	58.55	64.13	H^{α} 4.47, $H^{\beta a}$ 3.89, $H^{\beta b}$ 3.89
297 Ser	7.89	122.86	60.15	64.95	H^{α} 4.25, $H^{\beta a}$ 3.81, $H^{\beta b}$ 3.81

Table A4.1 PRK1 HR1c chemical shifts.

References

Abdul Azeez KR, Knapp S, Fernandes JMP, Klussmann E & Elkins JM (2014) The crystal structure of the RhoA–AKAP-Lbc DH–PH domain complex. *Biochem. J.* **464:** 231–239

Abramovitz A, Gutman M & Nachliel E (2012) Structural coupling between the Rho-insert domain of Cdc42 and the geranylgeranyl binding site of RhoGDI. *Biochemistry* **51:** 715–723

Aleku M, Schulz P, Keil O, Santel A, Schaeper U, Dieckhoff B, Janke O, Endruschat J, Durieux B, Röder N, Löffler K, Lange C, Fechtner M, Möpert K, Fisch G, Dames S, Arnold W, Jochims K, Giese K, Wiedenmann B, Scholz A & Kaufmann J (2008) Atu027, a liposomal small interfering RNA formulation targeting protein kinase N3, inhibits cancer progression. *Cancer Res.* **68:** 9788–9798

Amano M, Hamaguchi T, Shohag MH, Kozawa K, Kato K, Zhang X, Yura Y, Matsuura Y, Kataoka C, Nishioka T & Kaibuchi K (2015) Kinase-interacting substrate screening is a novel method to identify kinase substrates. *J. Cell Biol.* **209:** 895–912

Amano M, Mukai H, Ono Y, Chihara K, Matsui T, Hamajima Y, Okawa K, Iwamatsu A & Kaibuchi K (1996) Identification of a putative target for Rho as the serine-threonine kinase protein kinase N. *Science* **271:** 648–650

Amin E, Jaiswal M, Derewenda U, Reis K, Nouri K, Koessmeier KT, Aspenström P, Somlyo AV, Dvorsky R & Ahmadian MR (2016) Deciphering the molecular and functional basis of RHOGAP family proteins: A systematic approach toward selective inactivation of RHO family proteins. *J. Biol. Chem.* **291:** 20353–20371

Amor JC, Harrison DH, Kahn RA & Ringe D (1994) Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature* **372:** 704–708

Araki S, Kikuchi A, Hata Y, Isomura M & Takai Y (1990) Regulation of reversible binding of *smg* p25A, a *ras* p21-like GTP-binding protein, to synaptic plasma membranes and vesicles by its specific regulatory protein, GDP dissociation inhibitor. *J. Biol. Chem.* **265:**

13007–13015

Arencibia JM, Pastor-Flores D, Bauer AF, Schulze JO & Biondi RM (2013) AGC protein kinases: From structural mechanism of regulation to allosteric drug development for the treatment of human diseases. *Biochim. Biophys. Acta* **1834:** 1302–1321

Aspenström P, Ruusala A & Pacholsky D (2007) Taking Rho GTPases to the next level: The cellular functions of atypical Rho GTPases. *Exp. Cell Res.* **313:** 3673–3679

Balbo A & Schuck P (2005) Analytical ultracentrifugation in the study of protein self-association and heterogeneous protein-protein interactions. *Protein-Protein Interactions: A Molecular Cloning Manual. Edn 2. Edited by Golemis EA, Adams PD. Cold Spring Harbor Laboratory Press* 253–277

Balendran A, Biondi RM, Cheung PCF, Casamayor A, Deak M & Alessi DR (2000) A 3-phosphoinositide-dependent protein kinase-1 (PDK1) docking site is required for the phosphorylation of protein kinase Cζ (PKCζ) and PKC-related kinase 2 by PDK1. *J. Biol. Chem.* **275:** 20806–20813

Banno Y (2002) Regulation and possible role of mammalian phospholipase D in cellular functions. *J. Biochem.* **131:** 301–306

Bartsch JW, Mukai H, Takahashi N, Ronsiek M, Fuchs S, Jockusch H & Ono Y (1998) The protein kinase N (PKN) gene *PRKCL1/Prkcl1* maps to human chromosome 19p12-p13.1 and mouse chromosome 8 with close linkage to the myodystrophy (*myd*) mutation. *Genomics* **49:** 129–132

Bauer AF, Sonzogni S, Meyer L, Zeuzem S, Piiper A, Biondi RM & Neimanis S (2012) Regulation of protein kinase C-related protein kinase 2 (PRK2) by an intermolecular PRK2-PRK2 interaction mediated by its N-terminal domain. *J. Biol. Chem.* **287:** 20590–20602

Baumeister MA, Rossman KL, Sondek J & Lemmon MA (2006) The Dbs PH domain contributes independently to membrane targeting and regulation of guanine nucleotide-exchange activity. *Biochem. J.* **400:** 563–572

Berman HM, Kleywegt GJ, Nakamura H & Markley JL (2012) The Protein Data Bank at 40: Reflecting on the past to prepare for the future. *Structure* **20:** 391–396

Blumenstein L & Ahmadian MR (2004) Models of the cooperative mechanism for Rho effector recognition. *J. Biol. Chem.* **279:** 53419–53426

Bos J, Rehmann H & Wittinghofer A (2007) GEFs and GAPs: Critical elements in the control of small G proteins. *Cell* **129:** 865–877

Bourguignon LYW, Gilad E, Peyrollier K, Brightman A & Swanson RA (2007) Hyaluronan-CD44 interaction stimulates Rac1 signaling and PKNγ kinase activation leading to cytoskeleton function and cell migration in astrocytes. *J. Neurochem.* **101:** 1002–1017

Bourne HR, Sanders DA & McCormick F (1991) The GTPase superfamily: conserved structure and molecular mechanism. *Nature.* **349:** 117–127

Braverman LE & Quilliam LA (1999) Identification of Grb4/Nckβ, a Src homology 2 and 3 domain-containing adapter protein having similar binding and biological properties to Nck. *J. Biol. Chem.* **274:** 5542–5549

Brunet N, Morin A & Olofsson B (2002) RhoGDI-3 regulates RhoG and targets this protein to the Golgi complex through its unique N-terminal domain. *Traffic* **3:** 342–358

Brünger AT, Adams PD, Clore GM, Delano WL, Gros P, Grossekunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T & Warren GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54:** 905–921

Bustelo XR (2018) RHO GTPases in cancer: known facts, open questions, and therapeutic challenges. *Biochemical Society Transactions* **46:** 741–760

Calautti E, Grossi M, Mammucari C, Aoyama Y, Pirro M, Ono Y, Li J & Dotto GP (2002) Fyn tyrosine kinase is a downstream mediator of Rho/PRK2 function in keratinocyte cell-cell adhesion. *J. Cell Biol.* **156:** 137–148

Cann JR (1970) Interacting Macromolecules. *Academic Press, New York*

Cao J, Burke JE & Dennis EA (2013) Using hydrogen/deuterium exchange mass spectrometry to define the specific interactions of the phospholipase A_2 superfamily with lipid substrates, inhibitors, and membranes. *J. Biol. Chem.* **288:** 1806–1813

Casey PJ, Solski PA, Der CJ & Buss JE (1989) p21ras is modified by a farnesyl isoprenoid. *Proc. Natl. Acad. Sci. USA* **86:** 8323–8327

Cavanagh J, Fairbrother W, Palmer A, Rance M & Skelton N (1998) Protein NMR spectroscopy (Second Edition) *Academic Press*

Chamberlain P, Delker S, Pagarigan B, Mahmoudi A, Jackson P, Abbasian M, Muir J, Raheja N & Cathers B (2014) Crystal structures of PRK1 in complex with the clinical compounds lestaurtinib and tofacitinib reveal ligand induced conformational changes. *PLoS One* **9:** e103638

Chang F, Lemmon C, Lietha D, Eck M & Romer L (2011) Tyrosine phosphorylation of Rac1: A role in regulation of cell spreading. *PLoS One* **6:** e28587

Cheng Y, Zhu Y, Xu J, Yang M, Chen P, Xu W, Zhao J, Geng L & Gong S (2018) PKN2 in colon cancer cells inhibits M2 phenotype polarization of tumor-associated macrophages via regulating DUSP6-Erk1/2 pathway. *Mol. Cancer* **17:** 1–16

Cherfils J & Zeghouf M (2013) Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol. Rev.* **93:** 269–309

Chou CY, Hsieh YH & Chang GG (2011) Applications of analytical ultracentrifugation to protein size-and-shape distribution and structure-and-function analyses. *Methods* **54:** 76–82

Cole JL, Lary JW, Moody T & Laue TM (2008) Analytical ultracentrifugation: Sedimentation velocity and sedimentation equilibrium *Methods Cell Biol* **84:** 143–179

Collazos A, Michael N, Whelan RDH, Kelly G, Mellor H, Pang LCH, Totty N & Parker PJ (2011) Site recognition and substrate screens for PKN family proteins. *Biochem. J.* **438:** 535–543

Corbalán-García S & Gómez-Fernández JC (2010) The C2 domains of classical and novel PKCs

as versatile decoders of membrane signals. *BioFactors* **36:** 1–7

Côté JF & Vuori K (2007) GEF what? Dock180 and related proteins help Rac to polarize cells in new ways. *Trends Cell Biol.* **17:** 383–393

Crespo P, Schuebel KE, Ostrom AA, Gutkind JS & Bustelo XR (1997) Phosphotyrosine-dependent activation of Rac-1 GDP/GTP exchange by the *vav* proto-oncogene product. *Nature* **385:** 169–172

Cryns VL, Byun Y, Rana A, Mellor H, Lustig KD, Ghanem L, Parker PJ, Kirschner MW & Yuan J (1997) Specific proteolysis of the kinase protein kinase C-related kinase 2 by Caspase-3 during apoptosis. *J. Biol. Chem.* **272:** 29449–29453

D'Avino PP (2017) Citron kinase – renaissance of a neglected mitotic kinase. *J. Cell Sci.* **130:** 1701–1708

Danno S, Kubouchi K, Mehruba M, Abe M, Natsume R, Sakimura K, Eguchi S, Oka M, Hirashima M, Yasuda H & Mukai H (2017) PKN2 is essential for mouse embryonic development and proliferation of mouse fibroblasts. *Genes to Cells* **22:** 220–236

Delorme-Walker VD, Peterson JR, Chernoff J, Waterman CM, Danuser G, DerMardirossian C & Bokoch GM (2011) Pak1 regulates focal adhesion strength, myosin IIA distribution, and actin dynamics to optimize cell migration. *J. Cell Biol.* **193:** 1289–1303

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM & Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36:** 465–469

Dettori R, Sonzogni S, Meyer L, Lopez-Garcia LA, Morrice NA, Zeuzem S, Engel M, Piiper A, Neimanis S, Frödin M & Biondi RM (2009) Regulation of the interaction between protein kinase C-related protein kinase 2 (PRK2) and its upstream kinase, 3-phosphoinositide-dependent protein kinase 1 (PDK1). *J. Biol. Chem.* **284:** 30318–30327

Dong LQ, Landa LR, Wick MJ, Zhu L, Mukai H, Ono Y & Liu F (2000) Phosphorylation of protein kinase N by phosphoinositide-dependent protein kinase-1 mediates insulin signals to the actin cytoskeleton. *Proc. Natl. Acad. Sci. USA* **97:** 5089–5094

Dovas A, Choi Y, Yoneda A, Multhaupt HAB, Kwon SH, Kang D, Oh ES & Couchman JR (2010) Serine 34 phosphorylation of Rho guanine dissociation inhibitor (RhoGDIα) links signaling from conventional protein kinase C to RhoGTPase in cell adhesion. *J. Biol. Chem.* **285:** 23296–23308

Drozdetskiy A, Cole C, Procter J & Barton GJ (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* **43:** 389–394

Du & Prendergast (1999) Geranylgeranylated RhoB mediates suppression of human tumor cell growth by farnesyltransferase inhibitors. *Cancer Res.* **59:** 5492–5496

Dvorsky R, Blumenstein L, Vetter IR & Ahmadian MR (2004) Structural insights into the interaction of ROCKI with the switch regions of RhoA. *J. Biol. Chem.* **279:** 7098–7104

Falk MD, Liu W, Bolanos B, Unsal-Kacmaz K, Klippel A, Grant S, Brooun A & Timofeevski S (2014) Enzyme kinetics and distinct modulation of the protein kinase N family of kinases by lipid activators and small molecule inhibitors. *Biosci. Rep.* **34:** 93–106

Fekry M, Alshokry W, Grela P, Tchórzewski M, Ahlgren EC, Söderberg CA, Gakh O, Isaya G & Al-Karadaghi S (2017) SAXS and stability studies of iron-induced oligomers of bacterial frataxin CyaY. *PLoS One* **12:** e0184961

Fernandez-Borja M, Janssen L, Verwoerd D, Hordijk P & Neefjes J (2005) RhoB regulates endosome transport by promoting actin assembly on endosomal membranes through Dia1. *J. Cell Sci.* **118:** 2661–2670

Ferreira T, Prudêncio P & Martinho RG (2014) *Drosophila* protein kinase N (Pkn) is a negative regulator of actin-myosin activity during oogenesis. *Dev. Biol.* **394:** 277–291

Flynn P, Mellor H, Palmer R, Panayotou G, Parker PJ (1998) Multiple Interactions of PRK1 with RhoA. Functional Assignment of the HR1 repeat motif. *J. Biol. Chem.* **273:** 2698–2705

Flynn P, Mellor H, Casamassima A & Parker PJ (2000) Rho GTPase control of protein kinase C-related protein kinase activation by 3-phosphoinositide-dependent protein kinase. *J. Biol. Chem.* **275:** 11064–11070

Forget MA, Desrosiers RR, Gingras D & Béliveau R (2002) Phosphorylation states of Cdc42 and RhoA regulate their interactions with Rho GDP dissociation inhibitor and their extraction from biological membranes. *Biochem. J.* **361:** 243–254

Foster R, Hu KQ, Lu Y, Nolan KM, Thissen J & Settleman J (1996) Identification of a novel human Rho protein with unusual properties: GTPase deficiency and in vivo farnesylation. *Mol. Cell. Biol.* **16:** 2689–2699

Fukumoto Y, Kaibuchi K, Hori Y, Fujioka H, Araki S, Ueda T, Kikuchi A & Takai Y (1990) Molecular cloning and characterization of a novel type of regulatory protein (GDI) for the rho proteins, ras p21-like small GTP-binding proteins. *Oncogene* **5:** 1321–1328

Galgano MT, Conaway M, Spencer AM, Paschal BM & Frierson Jr. HF (2009) PRK1 distribution in normal tissues and carcinomas: overexpression and activation in ovarian serous carcinoma. *Hum. Pathol.* **40:** 1434–1440

Gampel A, Parker PJ & Mellor H (1999) Regulation of epidermal growth factor receptor traffic by the small GTPase RhoB. *Curr. Biol.* **9:** 955–958

Gao Q, Kumar A, Srinivasan S, Singh L, Mukai H, Ono Y, Wazer DE & Band V (2000) PKN binds and phosphorylates human papillomavirus E6 oncoprotein. *J. Biol. Chem.* **275:** 14824–14830

Gasman S, Kalaidzidis Y & Zerial M (2003) RhoD regulates endosome dynamics through Diaphanous-related Formin and Src tyrosine kinase. *Nat. Cell Biol.* **5:** 195–204

Gavard J & Gutkind JS (2008) Protein kinase C-related kinase and ROCK are required for thrombin-induced endothelial cell permeability downstream from $G\alpha_{12/13}$ and $G\alpha_{11/q}$. *J. Biol. Chem.* **283:** 29888–29896

Geyer M, Schweins T, Herrmann C, Prisner T, Wittinghofer A & Kalbitzer HR (1996) Conformational transitions in $p21^{ras}$ and in its complexes with the effector protein Raf-RBD and the GTPase activating protein GAP. *Biochemistry* **35:** 10308–10320

Gizachew D & Oswald R (2006) NMR structural studies of the myristoylated N-terminus of ADP ribosylation factor 6 (Arf6). *FEBS Lett.* **580:** 4296–4301

Goldberg J (1998) Structural basis for activation of ARF GTPase: Mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell* **95:** 237–248

Gosser YQ, Nomanbhoy TK, Aghazadeh B, Manor D, Combs C, Cerione RA & Rosen MK (1997) C-terminal binding domain of Rho GDP-dissociation inhibitor directs N- terminal inhibitory peptide to GTPases. *Nature* **387:** 814–819

Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1:** 2527–2535

Grizot S, Fauré J, Fieschi F, Vignais PV, Dagher MC & Pebay-Peyroula E (2001) Crystal structure of the Rac1–RhoGDI complex involved in NADPH oxidase activation. *Biochemistry* **40:** 10007–10013

Gross C, Heumann R & Erdmann KS (2001) The protein kinase C-related kinase PRK2 interacts with the protein tyrosine phosphatase PTP-BL via a novel PDZ domain binding motif. *FEBS Lett.* **496:** 101–104

Güldenhaupt J, Adigüzel Y, Kuhlmann J, Waldmann H, Kötting C & Gerwert K (2008) Secondary structure of lipidated Ras bound to a lipid bilayer. *FEBS J.* **275:** 5910–5918

Haga RB & Ridley AJ (2016) Rho GTPases: Regulation and roles in cancer cell biology. *Small GTPases* **7:** 207–221

Hakoshima T, Shimizu T & Maesaki R (2003) Structural basis of the Rho GTPase signaling. *J. Biochem.* **134:** 327–331

Han SH, Kim SJ, Kim EJ, Kim TE, Moon JS, Kim GW, Lee SH, Cho K, Yoo JS, Son WS, Rhee JK, Han SH & Oh JW (2014) Phosphorylation of hepatitis C virus RNA polymerases Ser29 and Ser42 by protein kinase C-related kinase 2 regulates viral RNA replication. *J. Virol.* **88:** 11240–11252

Hanzal-Bayer M, Renault L, Roversi P, Wittinghofer A & Hillig RC (2002) The complex of Arl2-GTP and PDEδ: from structure to function. *EMBO J.* **21:** 2095–2106

Harrison BC, Huynh K, Lundgaard GL, Helmke SM, Perryman MB & McKinsey TA (2010)

Protein kinase C-related kinase targets nuclear localization signals in a subset of class IIa histone deacetylases. *FEBS Lett.* **584:** 1103–1110

Hirshberg M, Stockley RW, Dodson G & Webb MR (1997) The crystal structure of human rac1, a member of the rho-family complexed with a GTP analogue. *Nat. Struct. Biol.* **4:** 147–152

Ho HYH, Rohatgi R, Lebensohn AM, Ma L, Li J, Gygi SP & Kirschner MW (2004) Toca-1 mediates Cdc42-dependent actin nucleation by activating the N-WASP-WIP complex. *Cell* **118:** 203–216

Hodge RG & Ridley AJ (2016) Regulating Rho GTPases and their regulators. *Nat. Rev. Mol. Cell Biol.* **17:** 496–510

Höfling S, Scharnert J, Cromme C, Bertrand J, Pap T, Schmidt MA & Rüter C (2014) Manipulation of pro-inflammatory cytokine production by the bacterial cell-penetrating effector protein YopM is independent of its interaction with host cell kinases RSK1 and PRK2. *Virulence* **5:** 761–771

Horiuchi H, Kawata M, Katayama M, Yoshida Y, Musha T, Ando S & Takai Y (1991) A novel prenyltransferase for a small GTP-binding protein having a C-terminal Cys-Ala-Cys structure. *J. Biol. Chem.* **266:** 16981–16984

Hu SH, Whitten AE, King GJ, Jones A, Rowland AF, James DE & Martin JL (2012) The weak complex between RhoGAP protein ARHGAP22 and signal regulatory protein 14-3-3 has 1:2 stoichiometry and a single peptide binding mode. *PLoS One* **7:** e41731

Hubbard S & Thornton J (1993) NACCESS, Computer Program. Department of Biochemistry and Molecular Biology, University College London.

Hutchinson CL, Lowe PN, McLaughlin SH, Mott HR & Owen D (2011) Mutational analysis reveals a single binding interface between RhoA and its effector, PRK1. *Biochemistry* **50:** 2860–2869

Hutchinson CL, Lowe PN, McLaughlin SH, Mott HR & Owen D (2013) Differential binding of RhoA, RhoB, and RhoC to protein kinase C-related kinase (PRK) isoforms PRK1, PRK2,

and PRK3: PRKs have the highest affinity for RhoB. *Biochemistry* **52:** 7999–8011

Iden S & Collard JG (2008) Crosstalk between small GTPases and polarity proteins in cell polarization. *Nat. Rev. Mol. Cell Biol.* **9:** 846–859

Isagawa T, Mukai H, Oishi K, Taniguchi T, Hasegawa H, Kawamata T, Tanaka C & Ono Y (2000) Dual effects of PKNα and protein kinase C on phosphorylation of tau protein by glycogen synthase kinase-3β. *Biochem. Biophys. Res. Commun.* **273:** 209–212

Ito H, Morishita R & Nagata KI (2018) Functions of Rhotekin, an effector of Rho GTPase, and its binding partners in mammals. *Int. J. Mol. Sci.* **19:** 2121

Ito Y, Yamasaki K, Iwahara J, Terada T, Kamiya A, Shirouzu M, Muto Y, Kawai G, Yokoyama S, Laue ED, Wälchli M, Shibata T, Nishimura S & Miyazawa T (1997) Regional polysterism in the GTP-bound form of the human c-Ha-Ras protein. *Biochemistry* **36:** 9109–9119

Itoh K, Yoshioka K, Akedo H, Uehata M, Ishizaki T & Narumiya S (1999) An essential part for Rho-associated kinase in the transcellular invasion of tumor cells. *Nat. Med.* **5:** 221–225

Jacinto E, Loewith R, Schmidt A, Lin S, Rüegg MA, Hall A & Hall MN (2004) Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. *Nat. Cell Biol.* **6:** 1122–1128

Jaiswal M, Fansa EK, Dvorsky R & Ahmadian MR (2012) New insight into the molecular switch mechanism of human Rho family proteins: shifting a paradigm. *Biol. Chem.* **394:** 89–95

Ju JA & Gilkes DM (2018) RhoB: Team oncogene or team tumor suppressor? *Genes* **9:** 1–18

Kajimoto K, Shao D, Takagi H, Maceri G, Zablocki D, Mukai H, Ono Y & Sadoshima J (2011) Hypotonic swelling-induced activation of PKN1 mediates cell survival in cardiac myocytes. *Am. J. Physiol. Heart Circ. Physiol.* **300:** 191-200

Karassek S, Berghaus C, Schwarten M, Goemans CG, Ohse N, Kock G, Jockers K, Neumann S, Gottfried S, Herrmann C, Heumann R & Stoll R (2010) Ras homolog enriched in brain (Rheb) enhances apoptotic signaling. *J. Biol. Chem.* **285:** 33979–33991

Katayama M, Kawata M, Yoshida Y, Horiuchi H, Yamamoto T, Matsuura Y & Takai Y (1991) The posttranslationally modified C-terminal structure of bovine aortic smooth muscle *rho*A p21. *J. Biol. Chem.* **266:** 12639–12645

Kawamata T, Taniguchi T, Mukai H, Kitagawa M, Hashimoto T, Maeda K, Ono Y & Tanaka C (1998) A protein kinase, PKN, accumulates in Alzheimer neurofibrillary tangles and associated endoplasmic reticulum-derived vesicles and phosphorylates tau protein. *J. Neurosci.* **18:** 7402–7410

Kay LE, Ikura M, Tschudin R & Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **89:** 496–514

Keeler J (2005) Understanding NMR Spectroscopy. *Wiley, Chichester*

Kelly SM, Jess TJ & Price NC (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta* **1751:** 119–139

Keszei AFA, Tang X, McCormick C, Zeqiraj E, Rohde JR, Tyers M & Sicheri F (2014) Structure of an SspH1-PKN1 complex reveals the basis for host substrate recognition and mechanism of activation for a bacterial E3 ubiquitin ligase. *Mol. Cell. Biol.* **34:** 362–373

Khwaja A, Sharpe CC, Noor M & Hendry BM (2006) The role of geranylgeranylated proteins in human mesangial cell proliferation. *Kidney Int.* **70:** 1296–1304

Kim MG, Moon JS, Kim EJ, Lee SH & Oh JW (2012) Destabilization of PDK1 by Hsp90 inactivation suppresses hepatitis C virus replication through inhibition of PRK2-mediated viral RNA polymerase phosphorylation. *Biochem. Biophys. Res. Commun.* **421:** 112–118

Kitagawa M, Mukai H, Shibata H & Ono Y (1995) Purification and characterization of a fatty acid-activated protein kinase (PKN) from rat testis. *Biochem. J* **310:** 657–664

Kitagawa M, Shibata H, Toshimori M, Mukai H & Ono Y (1996) The role of the unique motifs in the amino-terminal region of PKN on its enzymatic activity. *Biochem. Biophys. Res. Commun.* **220:** 963–968

Kleckner IR & Foster MP (2010) An introduction to NMR-based approaches for measuring

protein dynamics. *Biochim. Biophys. Acta* **1814:** 942–968

Kobashigawa Y, Kumeta H, Kanoh D & Inagaki F (2009) The NMR structure of the TC10- and Cdc42-interacting domain of CIP4. *J. Biomol. NMR* **44:** 113–118

Kozma R, Ahmed S, Best A & Lim L (1995) The Ras-related protein Cdc42Hs and bradykinin promote formation of peripheral actin microspikes and filopodia in Swiss 3T3 fibroblasts. *Mol. Cell. Biol.* **15:** 1942–1952

Kurella VB, Richard JM, Parke CL, LeCour Jr. LF, Bellamy HD & Worthylake DK (2009) Crystal structure of the GTPase-activating protein-related domain from IQGAP1. *J. Biol. Chem.* **284:** 14857–14865

Kwon T, Kwon DY, Chun J, Kim JH & Kang SS (2000) Akt protein kinase inhibits Rac1-GTP binding through phosphorylation at serine 71 of Rac1. *J. Biol. Chem.* **275:** 423–428

Lachmann S, Jevons A, De Rycker M, Casamassima A, Radtke S, Collazos A & Parker PJ (2011) Regulatory domain selectivity in the cell-type specific PKN-dependence of cell migration. *PLoS One* **6:** e21732

LaFevre-Bernt M, Sicheri F, Pico A, Porter M, Kuriyan J & Miller WT (1998) Intramolecular regulatory interactions in the Src family kinase Hck probed by mutagenesis of a conserved tryptophan residue. *J. Biol. Chem.* **273:** 32129–32134

Lamarche N & Hall A (1994) GAPs for rho-related GTPases. *Trends Genet.* **10:** 436–440

Lamm O (1929) Die Differentialgleichung der Ultrazentrifugierung. *Ark. Mat. Astr. Fys.* **21B:** 1-4

Lancaster CA, Taylor-Harris PM, Self AJ, Brill S, van Erp HE & Hall A (1994) Characterization of rhoGAP. *J. Biol. Chem.* **269:** 1137–1142

Laue TM, Shah BD, Ridgeway TM & Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In A*nalytical Ultracentrifugation in Biochemistry and Polymer Science. Edited by Harding SE, Rowe AJ & Horton JC.* The Royal Society of Chemistry, Cambridge, UK 90-125.

Lebowitz J, Lewis MS & Schuck P (2002) Modern analytical ultracentrifugation in protein science : A tutorial review. *Protein Sci.* **11:** 2067–2079

LeCour Jr. L, Boyapati VK, Liu J, Li Z, Sacks DB & Worthylake DK (2016) The structural basis for Cdc42-induced dimerization of IQGAPs. *Structure* **24:** 1499–1508

Lee SJ, Hwang J, Jeong HJ, Yoo M, Go GY, Lee JR, Leem YE, Park JW, Seo DW, Kim YK, Hahn MJ, Han JW, Kang JS & Bae GU (2016) PKN2 and Cdo interact to activate AKT and promote myoblast differentiation. *Cell Death Dis.* **7:** e2431

Leenders F, Möpert K, Schmiedeknecht A, Santel A, Czauderna F, Aleku M, Penschuck S, Dames S, Sternberger M, Röhl T, Wellmann A, Arnold W, Giese K, Kaufmann J & Klippel A (2004) PKN3 is required for malignant prostate cell growth downstream of activated PI 3-kinase. *EMBO J.* **23:** 3303–3313

Lepvrier E, Doigneaux C, Moullintraffort L, Nazabal A & Garnier C (2014) Optimized protocol for protein macrocomplexes stabilization using the EDC, 1-Ethyl-3-(3-(dimethylamino)propyl)carbodiimide, zero-length cross-linker. *Anal. Chem.* **86:** 10524–10530

Leroux AE, Schulze JO & Biondi RM (2018) AGC kinases, mechanisms of regulation and innovative drug development. *Semin. Cancer Biol.* **48:** 1–17

Lim MA, Yang L, Zheng Y, Wu H, Dong LQ & Liu F (2004) Roles of PDK-1 and PKN in regulating cell migration and cortical actin formation of PTEN-knockout cells. *Oncogene* **23:** 9348–9358

Lim WG, Chen X, Liu JP, Tan BJ, Zhou S, Smith A, Lees N, Hou L, Gu F, Yu XY, Du Y, Smith D, Verma C, Liu K & Duan W (2008) The C-terminus of PRK2/PKNγ is required for optimal activation by RhoA in a GTP-dependent manner. *Arch. Biochem. Biophys.* **479:** 170–178

Lim WG, Tan BJ, Zhu Y, Zhou S, Armstrong JS, Li QT, Dong Q, Chan E, Smith D, Verma C, Tan SL & Duan W (2006) The very C-terminus of PRK1/PKN is essential for its activation by RhoA and downstream signaling. *Cell. Signal.* **18:** 1473–81

Lim WG, Zhu Y, Wang CH, Tan BJ, Armstrong JS, Dokland T, Yang H, Zhu YZ, Teo TS & Duan W

(2005) The last five amino acid residues at the C-terminus of PRK1/PKN is essential for full lipid responsiveness. *Cell. Signal.* **17:** 1084–1097

Linge JP, O'Donoghue SI & Nilges M (2001) Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol.* **339:** 71-90

Liu X, Wang H, Eberstadt M, Schnuchel A, Olejniczak ET, Meadows RP, Schkeryantz JM, Janowick DA, Harlan JE, Harris EAS, Staunton DE & Fesik SW (1998) NMR structure and mutagenesis of the N-terminal Dbl homology domain of the nucleotide exchange factor Trio. *Cell* **95:** 269–277

Lodge JK, Johnson RL, Weinberg RA & Gordon JI (1994) Comparison of myristoyl-CoA: protein N-myristoyltransferases from three pathogenic fungi: *Cryptococcus neoformans*, *Histoplasma capsulatum*, and *Candida albicans*. *J. Biol. Chem.* **269:** 2996–3009

López-Alonso JP, Diez-García F, Font J, Ribó M, Vilanova M, Scholtz JM, González C, Vottariello F, Gotte G, Libonati M & Laurents DV (2009) Carbodiimide EDC induces cross-links that stabilize RNase A C-dimer against dissociation: EDC adducts can affect protein net charge, conformation, and activity. *Bioconjug. Chem.* **20:** 1459–1473

Lu Y & Settleman J (1999) The *Drosophila* Pkn protein kinase is a Rho/Rac effector target required for dorsal closure during embryogenesis. *Genes Dev.* **13:** 1168–1180

Maesaki R, Ihara K, Shimizu T, Kuroda S, Kaibuchi K & Hakoshima T (1999) The structural basis of Rho effector recognition revealed by the crystal structure of human RhoA complexed with the effector domain of PKN/PRK1. *Mol. Cell* **4:** 793–803

Makyio H, Ohgi M, Takei T, Takahashi S, Takatsu H, Katoh Y, Hanai A, Ueda T, Kanaho Y, Xie Y, Shin HW, Kamikubo H, Kataoka M, Kawasaki M, Kato R, Wakatsuki S & Nakayama K (2012) Structural basis for Arf6-MKLP1 complex formation on the Flemming body responsible for cytokinesis. *EMBO J.* **31:** 2590–2603

Manser C, Stevenson A, Banner S, Davies J, Tudor EL, Ono Y, Leigh PN, McLoughlin DM, Shaw CE & Miller CCJ (2008) Deregulation of PKN1 activity disrupts neurofilament organisation and axonal transport. *FEBS Lett.* **582:** 2303–2308

Marcos-Ramiro B, García-Weber D, Barroso S, Feito J, Ortega MC, Cernuda-Morollón E, Reglero-Real N, Fernández-Martín L, Durán MC, Alonso MA, Correas I, Cox S, Ridley AJ & Millán J (2016) RhoB controls endothelial barrier recovery by inhibiting Rac1 trafficking to the cell border. *J. Cell Biol.* **213:** 385–402

Markley JL, Ulrich EL, Berman HM, Henrick K, Nakamura H & Akutsu H (2008) BioMagResBank (BMRB) as a partner in the Worldwide Protein Data Bank (wwPDB): new policies affecting biomolecular NMR depositions. *J. Biomol. NMR* **40:** 153–155

Mashud R, Nomachi A, Hayakawa A, Kubouchi K, Danno S, Hirata T, Matsuo K, Nakayama T, Satoh R, Sugiura R, Abe M, Sakimura K, Wakana S, Ohsaki H, Kamoshida S & Mukai H (2017) Impaired lymphocyte trafficking in mice deficient in the kinase activity of PKN1. *Sci. Rep.* **7:** 1–14

Matsuzawa K, Kosako H, Inagaki N, Shibata H, Mukai H, Ono Y, Amano M, Kaibuchi K, Matsuura Y, Azuma I & Inagaki M (1997) Domain-specific phosphorylation of vimentin and glial fibrillary acidic protein by PKN. *Biochem. Biophys. Res. Commun.* **234:** 621–625

Mavaddat N, Mason DW, Atkinson PD, Evans EJ, Gilbert RJC, Stuart DI, Fennelly JA, Barclay AN, Davis SJ & Brown MH (2000) Signaling lymphocytic activation molecule (CDw150) is homophilic but self-associates with very low affinity. *J. Biol. Chem.* **275:** 28100–28109

McCarty OJT, Larson MK, Auger JM, Kalia N, Atkinson BT, Pearce AC, Ruf S, Henderson RB, Tybulewicz VLJ, Machesky LM & Watson SP (2005) Rac1 is essential for platelet lamellipodia formation and aggregate stability under flow. *J. Biol. Chem.* **280:** 39474–39484

McPhee JB, Mena P & Bliska JB (2010) Delineation of regions of the *Yersinia* YopM protein required for interaction with the RSK1 and PRK2 host kinases and their requirement for interleukin-10 production and virulence. *Infect. Immun.* **78:** 3529–3539

Mellor H, Flynn P, Nobes CD, Hall A & Parker PJ (1998) PRK1 is targeted to endosomes by the small GTPase, RhoB. *J. Biol. Chem.* **273:** 4811–4814

Metzger E, Müller JM, Ferrari S, Buettner R & Schüle R (2003) A novel inducible transactivation domain in the androgen receptor: implications for PRK in prostate

cancer. *EMBO J.* **22:** 270–280

Metzger E, Yin N, Wissmann M, Kunowska N, Fischer K, Friedrichs N, Patnaik D, Higgins JMG, Potier N, Scheidtmann KH, Buettner R & Schüle R (2008) Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation. *Nat. Cell Biol.* **10:** 53–60

Miki H, Yamaguchi H, Suetsugu S & Takenawa T (2000) IRSp53 is an essential intermediate between Rac and WAVE in the regulation of membrane ruffling. *Nature* **408:** 732–735

Misaki K, Mukai H, Yoshinaga C, Oishi K, Isagawa T, Takahashi M, Ohsumi K, Kishimoto T & Ono Y (2001) PKN delays mitotic timing by inhibition of Cdc25C: Possible involvement of PKN in the regulation of cell division. *Proc. Natl. Acad. Sci USA* **98:** 125–129

Mishra JP, Cohen D, Zamperone A, Nesic D, Muesch A & Stein M (2015) CagA of *Helicobacter pylori* interacts with and inhibits the serine-threonine kinase PRK2. *Cell. Microbiol.* **17:** 1670–1682

Modha R, Campbell LJ, Nietlispach D, Buhecha HR, Owen D & Mott HR (2008) The Rac1 polybasic region is required for interaction with its effector PRK1. *J. Biol. Chem.* **283:** 1492–500

Möpert K, Löffler K, Röder N, Kaufmann J & Santel A (2012) Depletion of protein kinase N3 (PKN3) impairs actin and adherens junctions dynamics and attenuates endothelial cell activation. *Eur. J. Cell Biol.* **91:** 694–705

Morrice NA, Gabrielli B, Kemp BE & Wettenhall REH (1994) A cardiolipin-activated protein kinase from rat liver structurally distinct from the protein kinases C. *J. Biol. Chem.* **269:** 20040–20046

Moskwa P, Paclet MH, Dagher MC & Ligeti E (2005) Autoinhibition of p50 Rho GTPase-activating protein (GAP) is released by prenylated small GTPases. *J. Biol. Chem.* **280:** 6716-6720

Mott HR & Owen D (2015) Structures of Ras superfamily effector complexes: What have we learnt in two decades? *Crit. Rev. Biochem. Mol. Biol.* **50:** 85–133

Movilla N, Dosil M, Zheng Y & Bustelo XR (2001) How Vav proteins discriminate the GTPases Rac1 and RhoA from Cdc42. *Oncogene* **20:** 8057–8065

Mukai H, Kitagawa M, Shibata H, Takanaga H, Mori K, Shimakawa M, Miyahara M, Hirao K & Ono Y (1994) Activation of PKN, a novel 120-kDa protein kinase with leucine zipper-like sequences, by unsaturated fatty acids and by limited proteolysis. *Biochem. Biophys. Res. Commun.* **204:** 348-356

Mukai H, Muramatsu A, Mashud R, Kubouchi K, Tsujimoto S, Hongu T, Kanaho Y, Tsubaki M, Nishida S, Shioi G, Danno S, Mehruba M, Satoh R & Sugiura R (2016) PKN3 is the major regulator of angiogenesis and tumor metastasis in mice. *Sci. Rep.* **6:** 18979

Mukai H & Ono Y (1994) A novel protein kinase with leucine zipper-like sequences: its catalytic domain is highly homologous to that of protein kinase C. *Biochem. Biophys. Res. Commun.* **199:** 897-904

Mukai H & Ono Y (2006) Purification and kinase assay of PKN. *Methods Enzymol.* **406:** 234–50

Mukai H, Toshimori M, Shibata H, Takanaga H, Kitagawa M, Miyahara M, Shimakawa M & Ono Y (1997) Interaction of PKN with α-actinin. *J. Biol. Chem.* **272:** 4740–4746

Murayama K, Shirouzu M, Kawasaki Y, Kato-Murayama M, Hanawa-Suetsugu K, Sakamoto A, Katsura Y, Suenaga A, Toyama M, Terada T, Taiji M, Akiyama T & Yokoyama S (2007) Crystal structure of the Rac activator, Asef, reveals its autoinhibitory mechanism. *J. Biol. Chem.* **282:** 4238–4242

Narumiya S, Tanji M & Ishizaki T (2009) Rho signaling, ROCK and mDia1, in transformation, metastasis and invasion. *Cancer Metastasis Rev.* **28:** 65–76

Nassar N, Horn G, Herrmann C, Scherer A, McCormick F & Wittinghofer A (1995) The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. *Nature* **375:** 554–560

Nilges M (1996) Structure calculation from NMR data. *Curr. Opin. Struct. Biol.* **6:** 617–623

Nilsson J, Weis K & Kjems J (2002) The C-terminal extension of the small GTPase Ran is

essential for defining the GDP-bound form. *J. Mol. Biol.* **318:** 583–593

Nobes CD & Hall A (1995) Rho, Rac, and Cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* **81:** 53–62

Nomanbhoy TK, Erickson JW & Cerione RA (1999) Kinetics of Cdc42 membrane extraction by Rho-GDI monitored by real-time fluorescence resonance energy transfer. *Biochemistry* **38:** 1744–1750

O'Sullivan AG, Mulvaney EP, Hyland PB & Kinsella BT (2015) Protein kinase C-related kinase 1 and 2 play an essential role in thromboxane-mediated neoplastic responses in prostate cancer. *Oncotarget* **6:** 26437-26456

O'Sullivan AG, Mulvaney EP & Kinsella BT (2017) Regulation of protein kinase C-related kinase (PRK) signalling by the TPα and TPβ isoforms of the human thromboxane A_2 receptor: Implications for thromboxane- and androgen- dependent neoplastic and epigenetic responses in prostate cancer. *Biochim. Biophys. Acta* **1863:** 838–856

Oishi K, Mukai H, Shibata H, Takahashi M & Ona Y (1999) Identification and characterization of PKNβ, a novel isoform of protein kinase PKN: Expression and arachidonic acid dependency are different from those of PKNα. *Biochem. Biophys. Res. Commun.* **261:** 808–14

Oishi K, Takahashi M, Mukai H, Banno Y, Nakashima S, Kanaho Y, Nozawa Y & Ono Y (2001) PKN regulates phospholipase D1 through direct interaction. *J. Biol. Chem.* **276:** 18096–18101

Olson MF, Ashworth A & Hall A (1995) An essential role for Rho, Rac, and Cdc42 GTPases in cell cycle progression through G1. *Science* **269:** 1270–1272

Ortega A, Amorós D & García de la Torre J (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101:** 892–898

Owen D, Lowe PN, Nietlispach D, Brosnan CE, Chirgadze DY, Parker PJ, Blundell TL & Mott

HR (2003) Molecular dissection of the interaction between the small G proteins Rac1 and RhoA and protein kinase C-related kinase 1 (PRK1). *J. Biol. Chem.* **278:** 50578–50587

Owen D, Campbell LJ, Littlefield K, Evetts KA, Li Z, Sacks DB, Lowe PN & Mott HR (2008) The IQGAP1-Rac1 and IQGAP1-Cdc42 interactions. *J. Biol. Chem.* **283:** 1692–1704

Palmer RH, Dekker LV, Woscholski R, Le Good JA, Gigg R & Parker PJ (1995) Activation of PRK1 by phosphatidylinositol 4,5-bisphosphate and phosphatidylinositol 3,4,5-trisphosphate. *J. Biol. Chem.* **270:** 22412–22416

Park YH, Wood G, Kastner DL & Chae JJ (2016) Pyrin inflammasome activation and RhoA signaling in the autoinflammatory diseases FMF and HIDS. *Nat. Immunol.* **17:** 914–921

Parrini MC, Lei M, Harrison SC & Mayer BJ (2002) Pak1 kinase homodimers are autoinhibited in *trans* and dissociated upon activation by Cdc42 and Rac1. *Mol. Cell* **9:** 73–83

Pasqualato S, Renault L & Cherfils J (2002) Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication. *EMBO Rep.* **3:** 1035–1041

Paul F, Zauber H, von Berg L, Rocks O, Daumke O & Selbach M (2017) Quantitative GTPase affinity purification identifies Rho family protein interaction partners. *Mol. Cell. Proteomics* **16:** 73–85

Pearce LR, Komander D & Alessi DR (2010) The nuts and bolts of AGC protein kinases. *Nat. Rev. Mol. Cell Biol.* **11:** 9–22

Pechlivanis M, Ringel R, Popkirova B & Kuhlmann J (2007) Prenylation of Ras facilitates hSOS1-promoted nucleotide exchange, upon Ras binding to the regulatory site. *Biochemistry* **46:** 5341-5348

Peng B, Morrice NA, Groenen LC & Wettenhall REH (1996) Phosphorylation events associated with different states of activation of a hepatic cardiolipin/protease-activated protein kinase. *J. Biol. Chem.* **271:** 32233–32240

Piotto M, Saudek V & Sklenár V (1992) Gradient-tailored excitation for single-quantum NMR

spectroscopy of aqueuous solutions. *J. Biomol. NMR* **2:** 661–665

Pollard TD (2010) A guide to simple and informative binding assays. *Mol. Biol. Cell* **21:** 4061–4067

Porter AP, Papaioannou A & Malliri A (2016) Deregulation of Rho GTPases in cancer. *Small GTPases* **7:** 123–138

Prakash P & Gorfe AA (2017) Membrane orientation dynamics of lipid-modified small GTPases. *Small GTPases* **8:** 129–138

Pylypenko O, Rak A, Durek T, Kushnir S, Dursina BE, Thomae NH, Constantinescu AT, Brunsveld L, Watzke A, Waldmann H, Goody RS & Alexandrov K (2006) Structure of doubly prenylated Ypt1:GDI complex and the mechanism of GDI-mediated Rab recycling. *EMBO J.* **25:** 13–23

Quétier I, Marshall JJT, Spencer-Dene B, Lachmann S, Casamassima A, Franco C, Escuin S, Worrall JT, Baskaran P, Rajeeve V, Howell M, Copp AJ, Stamp G, Rosewell I, Cutillas P, Gerhardt H, Parker PJ & Cameron AJM (2016) Knockout of the PKN family of Rho effector kinases reveals a non-redundant role for PKN2 in developmental mesoderm expansion. *Cell Rep.* **14:** 440–448

Quilliam LA, Lambert QT, Mickelson-Young LA, Westwick JK, Sparks AB, Kay BK, Jenkins NA, Gilbert DJ, Copeland NG & Der CJ (1996) Isolation of a NCK-associated kinase, PRK2, an SH3-binding protein and potential effector of Rho protein signaling. *J. Biol. Chem.* **271:** 28772–28776

Ralston G (1993) Introduction to analytical ultracentrifugation. *Beckman Instruments, Fullerton, CA*

Ranjbar B & Gill P (2009) Circular dichroism techniques: Biomolecular and nanostructural analyses - A review. *Chem. Biol. Drug Des.* **74:** 101–120

Reid T, Furuyashiki T, Ishizaki T, Watanabe G, Watanabe N, Fujisawa K, Morii N, Madaule P & Narumiya S (1996) Rhotekin, a new putative target for Rho bearing homology to a serine/threonine kinase, PKN, and rhophilin in the Rho-binding domain. *J. Biol. Chem.*

271: 13556–13560

Ridley AJ (2015) Rho GTPase signalling in cell migration. *Curr. Opin. Cell Biol.* **36:** 103–112

Ridley AJ, Paterson HF, Johnston CL, Diekmann D & Hall A (1992) The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* **70:** 401–410

Ridley AJ & Hall A (1992) The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **70:** 389–399

Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE & Nilges M (2007) ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23:** 381–382

Rittinger K, Walker PA, Eccleston JF, Nurmahomed K, Owen D, Laue E, Gamblin SJ & Smerdon SJ (1997) Crystal structure of a small G protein in complex with the GTPase-activating protein rhoGAP. *Nature* **388:** 693–697

Roberts PJ, Mitin N, Keller PJ, Chenette EJ, Madigan JP, Currin RO, Cox AD, Wilson O, Kirschmeier P & Channing JD (2008) Rho family GTPase modification and dependence on CAAX motif-signaled posttranslational modification *J. Biol. Chem.* **283:** 25150–25163

Rohatgi R, Ho HH & Kirschner MW (2000) Mechanism of N-WASP activation by CDC42 and phosphatidylinositol 4,5-bisphosphate. *J. Cell Biol.* **150:** 1299–1309

Rolli-Derkinderen M, Sauzeau V, Boyer L, Lemichez E, Baron C, Henrion D, Loirand G & Pacaud P (2005) Phosphorylation of serine 188 protects RhoA from ubiquitin/proteasome-mediated degradation in vascular smooth muscle cells. *Circ. Res.* **96:** 1152–1160

Rose R, Weyand M, Lammers M, Ishizaki T, Ahmadian MR & Wittinghofer A (2005) Structural and mechanistic insights into the interaction between Rho and mammalian Dia. *Nature* **435:** 513–518

Rosman KJR & Taylor PDP (1998) Isotopic compositions of the elements 1997 *J. Phys. Chem. Ref. Data* **27:** 1275–1287

Rule GS & Hitchens TK (2006) Fundamentals of protein NMR spectroscopy. *Springer*

Sandberg A, Luheshi LM, Söllvander S, Pereira de Barros T, Macao B, Knowles TPJ, Biverstål H, Lendel C, Ekholm-Petterson F, Dubnovitsky A, Lannfelt L, Dobson CM & Härd T (2010) Stabilization of neurotoxic Alzheimer amyloid-β oligomers by protein engineering. *Proc. Natl. Acad. Sci. USA* **107:** 15595–15600

Sandilands E & Frame MC (2008) Endosomal trafficking of Src tyrosine kinase. *Trends Cell Biol.* **18:** 322–329

Sasaki T, Kikuchi A, Araki S, Hata Y, Isomura M, Kuroda S & Takai Y (1990) Purification and characterization from bovine brain cytosol of a protein that inhibits the dissociation of GDP from and the subsequent binding of GTP to *smg* p25A, a *ras* p21-like GTP-binding protein. *J. Biol. Chem.* **265:** 2333-2337

Sass GL & Ostrow BD (2014) Disruption of the *protein kinase N* gene of *Drosophila melanogaster* results in the recessive *delorean* allele (pkn^{dln}) with a negative impact on wing morphogenesis. *Genes*|*Genomes*|*Genetics* **4:** 643–656

Schachman HK (1959) Ultracentrifugation in biochemistry. *Academic Press, New York*

Schafer WR, Kim R, Sterne R, Thorner J, Kim SH & Rine J (1989) Genetic and pharmacological suppression of oncogenic mutations in *RAS* genes of yeast and humans. *Science* **245:** 379–385

Scheffzek K, Klebe C, Fritz-Wolf K, Kabsch W & Wittinghofer A (1995) Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form. *Nature* **374:** 378–381

Schmidt A, Durgan J, Magalhaes A & Hall A (2007) Rho GTPases regulate PRK2/PKN2 to control entry into mitosis and exit from cytokinesis. *EMBO J.* **26:** 1624–1636

Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78:** 1606–1619

Schwarten M, Berghaus C, Heumann R & Stoll R (2007) Sequence-specific ^{1}H, ^{13}C, and ^{15}N backbone assignment of the activated 21 kDa GTPase rRheb. *Biomol. NMR Assign.* **1:** 105–108

Seabra MC, Goldstein JL, Südhof TC & Brown MS (1992) Rab geranylgeranyl transferase. A multisubunit enzyme that prenylates GTP-binding proteins terminating in Cys-X-Cys or Cys-Cys. *J. Biol. Chem.* **267:** 14497–14503

Sebbagh M, Renvoizé C, Hamelin J, Riché N, Bertoglio J & Bréard J (2001) Caspase-3-mediated cleavage of ROCK I induces MLC phosphorylation and apoptotic membrane blebbing. *Nat. Cell Biol.* **3:** 346–352

Shen Y & Bax A (2010) SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J. Biomol. NMR* **48:** 13–22

Shen Y & Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56:** 227–241

Shibata H, Oda H, Mukai H, Oishi K, Misaki K, Ohkubo H & Ono Y (1999) Interaction of PKN with a neuron-specific basic helix-loop-helix transcription factor, NDRF/NeuroD2. *Brain Res. Mol. Brain Res.* **74:** 126-134

Shibata H, Oishi K, Yamagiwa A, Matsumoto M, Mukai H & Ono Y (2001) PKNβ interacts with the SH3 domains of Graf and a novel Graf related protein, Graf2, which are GTPase activating proteins for Rho family. *J. Biochem.* **130:** 23–31

Shiga K, Takayama K, Futaki S, Hutti JE, Cantley LC, Ueki K, Ono Y & Mukai H (2010) Development of an intracellularly acting inhibitory peptide selective for PKN. *Biochem. J.* **425:** 445–453

Singh NK, Kundumani-Sridharan V, Kumar S, Verma SK, Kotla S, Mukai H, Heckle MR & Rao GN (2012) Protein kinase N1 Is a novel substrate of NFATc1-mediated cyclin D1-CDK6 activity and modulates vascular smooth muscle cell division and migration leading to inward blood vessel wall remodeling. *J. Biol. Chem.* **287:** 36291–36304

Slynko I, Schmidtkunz K, Rumpf T, Klaeger S, Heinzlmeir S, Najar A, Metzger E, Kuster B, Schüle R, Jung M & Sippl W (2016) Identification of highly potent protein kinase C-related kinase 1 inhibitors by virtual screening, binding free energy rescoring, and in vitro testing. *ChemMedChem* **11:** 2084–2094

Snyder DA, Chen Y, Denissova NG, Acton T, Aramini JM, Ciano M, Karlin R, Liu J, Manor P, Rajan PA, Rossi P, Swapna GVT, Xiao R, Rost B, Hunt J & Montelione GT (2005) Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are complementary methods for small protein structure determination. *J. Am. Chem. Soc.* **127:** 16505–16511

Snyder JT, Worthylake DK, Rossman KL, Betts L, Pruitt WM, Siderovski DP, Der CJ & Sondek J (2002) Structural basis for the selective activation of Rho GTPases by Dbl exchange factors. *Nat. Struct. Biol.* **9:** 468–475

Sophocleous G (2014) Structural and functional analysis of the HR1c domain of PRK3, a protein kinase C- related kinase. *MSci Thesis, University of Cambridge*

Spera S & Bax A (1991) Empirical correlation between protein backbone conformation of Cα and Cβ ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113:** 5490–5492

Sreerama N & Woody RW (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem.* **209:** 32-44

Standaert M, Bandyopadhyay G, Galloway L, Ono Y, Mukai H & Farese R (1998) Comparative effects of GTPγS and insulin on the activation of Rho, phosphatidylinositol 3-kinase, and protein kinase N in rat adipocytes. *J.Biol. Chem.* **273:** 7470–7477

Strumberg D, Schultheis B, Traugott U, Vank C, Santel A, Keil O, Giese K, Kaufmann J & Drevs J (2012) Phase I clinical development of Atu027, a siRNA formulation targeting PKN3 in patients with advanced solid tumors. *Int. J. Clin. Pharmacol. Ther.* **50:** 76–78

Sun W, Vincent S, Settleman J & Johnson GL (2000) MEK kinase 2 binds and activates protein kinase C-related kinase 2. Bifurcation of kinase regulatory pathways at the level of an MAPK kinase kinase. *J. Biol. Chem.* **275:** 24421–24428

Svergun DI (2007) Small-angle scattering studies of macromolecular solutions. *J. Appl. Cryst.* **40:** s10–s17

Swarthout JT, Lobo S, Farh L, Croke MR, Greentree WK, Deschenes RJ & Linder ME (2005)

DHHC9 and GCP16 constitute a human protein fatty acyltransferase with specificity for H- and N-Ras. *J. Biol. Chem.* **280:** 31141–31148

Symons M & Rusk N (2003) Control of vesicular trafficking by Rho GTPases. *Curr. Biol.* **13:** 409–418

Taguchi T & Misaki R (2011) Palmitoylation pilots Ras to recycling endosomes. *Small GTPases* **2:** 82–84

Takagi H, Hsu CP, Kajimoto K, Shao D, Yang Y, Maejima Y, Zhai P, Yehia G, Yamada C, Zablocki D & Sadoshima J (2010) Activation of PKN mediates survival of cardiac myocytes in the heart during ischemia/reperfusion. *Circ. Res.* **107:** 642–649

Takahashi M, Mukai H, Toshimori M, Miyamoto M & Ono Y (1998) Proteolytic activation of PKN by caspase-3 or related protease during apoptosis. *Proc. Natl. Acad. Sci. USA* **95:** 11566–11571

Tarricone C, Xiao B, Justin N, Walker PA, Rittinger K, Gamblin SJ & Smerdon SJ (2001) The structural basis of Arfaptin-mediated cross-talk between Rac and Arf signalling pathways. *Nature* **411:** 215–219

Thauerer B, zur Nedden S & Baier-Bitterlich G (2014) Protein kinase C-related kinase (PKN/PRK). Potential key-role for PKN1 in protection of hypoxic neurons. *Curr. Neuropharmacol.* **12:** 213–218

Thumkeo D, Watanabe S & Narumiya S (2013) Physiological roles of Rho and Rho effectors in mammals. *Eur. J. Cell Biol.* **92:** 303–315

Tong J, Li L, Ballermann B & Wang Z (2013) Phosphorylation of Rac1 T108 by extracellular signal-regulated kinase in response to epidermal growth factor: a novel mechanism to regulate Rac1 function. *Mol. Cell. Biol.* **33:** 4538–4551

Torbett NE, Casamassima A & Parker PJ (2003) Hyperosmotic-induced protein kinase N 1 activation in a vesicular compartment is dependent upon Rac1 and 3-phosphoinositide-dependent kinase 1. *J. Biol. Chem.* **278:** 32344–32351

Turner EC, Kavanagh DJ, Mulvaney EP, McLean C, Wikström K, Reid HM & Kinsella BT (2011)

Identification of an interaction between the TPα and TPβ isoforms of the human thromboxane A_2 receptor with protein kinase C-related kinase (PRK) 1: Implications for prostate cancer. *J. Biol. Chem.* **286:** 15440–15457

Ueyama T, Ren Y, Sakai N, Takahashi M, Ono Y, Kondoh T, Tamaki N & Saito N (2001) Generation of a constitutively active fragment of PKN in microglia/macrophages after middle cerebral artery occlusion in rats. *J. Neurochem.* **79:** 903–913

Unsal-Kacmaz K, Ragunathan S, Rosfjord E, Dann S, Upeslacis E, Grillo M, Hernandez R, Mack F & Klippel A (2012) The interaction of PKN3 with RhoC promotes malignant growth. *Mol. Oncol.* **6:** 284–98

Vega FM, Fruhwirth G, Ng T & Ridley AJ (2011) RhoA and RhoC have distinct roles in migration and invasion by acting through different targets. *J. Cell Biol.* **193:** 655–665

Vega FM & Ridley AJ (2007) SnapShot: Rho family GTPases. *Cell* **129:** 1430

Vidal C, Geny B, Melle J, Jandrot-Perrus M & Fontenay-Roupie M (2002) Cdc42/Rac1-dependent activation of the p21-activated kinase (PAK) regulates human platelet lamellipodia spreading: implication of the cortical-actin binding protein cortactin. *Blood* **100:** 4462–4469

Vincent S & Settleman J (1997) The PRK2 kinase is a potential effector target of both Rho and Rac GTPases and regulates actin cytoskeletal organization. *Mol. Cell Biol.* **17:** 2247-2256

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J & Laue ED (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59:** 687–696

Walker JE, Saraste M, Runswick MJ & Gay NJ (1982) Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymees and a common nucleotide binding fold. *EMBO J.* **1:** 945–951

Wallace SW, Magalhaes A & Hall A (2011) The Rho target PRK2 regulates apical junction formation in human bronchial epithelial cells. *Mol. Cell. Biol.* **31:** 81–91

Wang F, Zhan R, Chen L, Dai X, Wang W, Guo R, Li X, Li Z, Wang L, Huang S, Shen J, Li S & Cao C (2017) RhoA promotes epidermal stem cell proliferation via PKN1-cyclin D1 signaling. *PLoS One* **12:** e0172613

Wang M & Casey PJ (2016) Protein prenylation: unique fats make their mark on biology. *Nat. Rev. Mol. Cell Biol.* **17:** 110–122

Watanabe G, Saito Y, Madaule P, Ishizaki T, Fujisawa K, Morii N, Mukai H, Ono Y, Kakizuka A & Narumiya S (1996) Protein kinase N (PKN) and PKN-related protein rhophilin as targets of small GTPase Rho. *Science* **271:** 645–648

Watanabe T, Wang S & Kaibuchi K (2015) IQGAPs as key regulators of actin-cytoskeleton dynamics. *Cell Struct. Funct.* **40:** 69–77

Watson JR, Fox HM, Nietlispach D, Gallop JL, Owen D & Mott HR (2016) Investigation of the interaction between Cdc42 and its effector TOCA1. *J. Biol. Chem.* **291:** 13875–13890

Watson JR (2017) Structural and biochemical insight into the interactions of Cdc42 with TOCA1 and N-WASP. *PhD Thesis, University of Cambridge*

Wennerberg K, Rossman KL & Der CJ (2005) The Ras superfamily at a glance. *J. Cell Sci.* **118:** 843–846

Whitmore L & Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **89:** 392–400

Wikström K, Kavanagh DJ, Reid HM & Kinsella BT (2008) Differential regulation of RhoA-mediated signaling by the TPα and TPβ isoforms of the human thromboxane A_2 receptor: Independent modulation of TPα signaling by prostacyclin and nitric oxide. *Cell. Signal.* **20:** 1497–1512

Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73:** 1–16

Wittinghofer A & Vetter IR (2011) Structure-function relationships of the G Domain, a canonical switch motif. *Annu. Rev. Biochem.* **80:** 943–971

Worthylake DK, Rossman KL & Sondek J (2000) Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1. *Nature* **408:** 682–688

Wu WJ, Lin R, Cerione RA & Manor D (1998) Transformation activity of Cdc42 requires a region unique to Rho-related proteins. *Biochemistry* **273:** 16655–16658

Wu YW, Tan KT, Waldmann H, Goody RS & Alexandrov K (2007) Interaction analysis of prenylated Rab GTPase with Rab escort protein and GDP dissociation inhibitor explains the need for both regulators. *Proc. Natl. Acad. Sci. USA* **104:** 12294–12299

Yang CS, Melhuish TA, Spencer A, Ni L, Hao Y, Jividen K, Harris TE, Snow C, Frierson Jr. HF, Wotton D & Paschal BM (2017) The protein kinase C super-family member PKN is regulated by mTOR and influences differentiation during prostate cancer progression. *Prostate* **77:** 1452–1467

Yasui T, Sakakibara-Yada K, Nishimura T, Morita K, Tada S, Mosialos G, Kieff E & Kikutani H (2012) Protein kinase N1, a cell inhibitor of Akt kinase, has a central role in quality control of germinal center formation. *Proc. Natl. Acad. Sci. USA* **109:** 21022–21027

Yoshinaga C, Mukai H, Toshimori M, Miyamoto M & Ono Y (1999) Mutational analysis of the regulatory mechanism of PKN: The regulatory region of PKN contains an arachidonic acid-sensitive autoinhibitory domain. *J. Biochem.* **126:** 475–484

Yu W, Liu J, Morrice NA & Wettenhall REH (1997) Isolation and characterization of a structural homologue of human PRK2 from rat liver: Distinguishing substrate and lipid activator specificities. *J. Biol. Chem.* **272:** 10030–10034

Yu J, Zhang D, Liu J, Li J, Yu Y, Wu XR & Huang C (2012) RhoGDI SUMOylation at Lys-138 increases its binding activity to Rho GTPase and its inhibiting cancer cell motility. *J. Biol. Chem.* **287:** 13752–13760

Zegers MM & Friedl P (2014) Rho GTPases in collective cell migration. *Small GTPases* **5:** e983869

Zhu Y, Stolz DB, Guo F, Ross MA, Watkins SC, Tan BJ, Qi RZ, Manser E, Li QT, Bay BH, Teo TS & Duan W (2004) Signaling via a novel integral plasma membrane pool of a

serine/threonine protein kinase PRK1 in mammalian cells. *FASEB J.* **18:** 1722–1724

Zong H, Kaibuchi K & Quilliam LA (2001) The insert region of RhoA is essential for Rho kinase activation and cellular transformation. *Mol. Cell. Biol.* **21:** 5287–5298

zur Nedden S, Eith R, Schwarzer C, Zanetti L, Seitter H, Fresser F, Koschak A, Cameron AJM, Parker PJ, Baier G & Baier-Bitterlich G (2018) Protein kinase N1 critically regulates cerebellar development and long-term function. *J. Clin. Invest.* **128:** 2076–2088